

PABLO RODRÍGUEZ

PRÓLOGO DE SANDY PENTLAND, PROFESOR DEL MIT MEDIA LAB



INTELIGENCIA ARTIFICIAL

CÓMO CAMBIARÁ EL MUNDO
(Y TU VIDA)



En la actualidad, la recolección y análisis de datos ha adquirido un papel esencial en nuestras vidas, no sólo para otorgar una imagen certera de una situación concreta sino también para predecir actuaciones o posibilidades futuras. Los datos y la inteligencia artificial 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. Además, sus posibilidades son inmensas, los macrodatos y la IA pueden ayudarnos a predecir variables como los precios de mercado, las sequías, las migraciones o una futura escasez de alimentos.

En este libro, Pablo Rodríguez realiza un análisis de las dimensiones que los macrodatos y la IA están adquiriendo en nuestra sociedad y como están transformado aspectos como la seguridad, la salud, las finanzas o el propio aprendizaje.



Pablo Rodríguez

Inteligencia artificial

Cómo cambiará el mundo (y tu vida)

ePub r1.2

Titivillus 29.03.2020

Título original: *Inteligencia artificial*
Pablo Rodríguez, 2018
Traducción: Jorge Rizzo

Editor digital: Titivillus
ePub base r2.1



Índice de contenido

Inteligencia artificial

Prólogo

Introducción

1 El regreso al futuro de los macrodatos

¿Por qué deberías preocuparte?

El futuro empieza ahora

Un viaje al interior de los datos

Historia antigua de los datos

La emergencia de la estadística

Los primeros días del almacenaje moderno de datos

Los inicios de la inteligencia empresarial

La aparición de los grandes centros de datos

La aparición de internet

Las primeras ideas sobre macrodatos

La web 2.0 aumenta los volúmenes de datos

Se adopta el uso actual del término «big data» (macrodatos)

El poder de las redes de datos

Interconectando Datos: propiedades mágicas

Científicos de datos: los gurús de los macrodatos

2 El poder de las cosas y sus datos

La Internet de las Cosas

Cómo acabar con internet

Una distribución de datos más rápida

La IdC a la velocidad de la luz

La Internet de las Cosas y los datos

Seguridad para todos

3 Inteligencia artificial: humanos contra máquinas

IA

Algoritmos

Aprendizaje automático

Aprendizaje cognitivo

La IA te cambia la vida

¿Te da miedo la IA?

La IA y los macrodatos

4 La IA para el liderazgo y la creatividad

Liderazgo y fútbol: el colectivismo y el deporte

El fútbol como algoritmo de datos
Un laboratorio de datos en el campo
El ordenador en la cocina
El lenguaje de los sabores
La IA y la gastronomía

5 Apostarlo todo a la ciencia de datos y a la IA
Un paseo por La Boqueria
La IA y el turismo
La IA y la seguridad
La IA y las finanzas
La IA, el empleo y la movilidad

6 La IA en busca del bien global
Un día en Harvard
La IA y la salud mental
La salud física y la IA
La IA para el bien global

7 IA ética, privacidad y el alma de los datos
La utopía del ciberespacio
La privacidad de los datos
Un delicado equilibrio
Bancos de datos personales

Conclusión

Bibliografía

Agradecimientos

Autor

Para Uriel

Prólogo

Casi todos los problemas importantes del ser humano se pueden abordar usando datos y la inteligencia artificial (IA). Pero históricamente ha sido imposible, dada la falta de referencias y la falta de enlaces que permitan el análisis de los datos.

En la última década se ha registrado un cambio enorme en el mundo de la IA y los datos. Quizá el mayor avance de todos sea que ahora el 80 por ciento de la totalidad de los seres humanos tenemos acceso a un teléfono y todo el mundo está conectado digitalmente al resto del mundo. Hasta las comunidades más pobres tienen acceso a las telecomunicaciones. Y los registros de esas interacciones miden aspectos importantes de la sociedad humana que los utiliza, porque los datos que generan al conectarse entre ellos o con las redes de telecomunicaciones, nos aportan cifras de población, patrones de movimiento y patrones de interacción entre comunidades. Y como consecuencia, ahora disponemos de informaciones sobre las condiciones de vida de los seres humanos en casi todos los rincones del mundo, y sobre los enlaces de conexión usados para obtener esos datos.

Esta nueva infraestructura de comunicaciones nos ha aportado posibilidades absolutamente impresionantes para utilizar la IA. Los registros nos dan todo tipo de información sobre nuestro entorno inmediato y podemos interactuar con ellos con la IA de un modo que hace diez años habríamos considerado una locura o algo carísimo. ¿Quieres saber cómo está el tráfico de camino a casa? ¿Qué restaurantes les gustan a tus amigos? ¿Qué tiendas del barrio están abiertas? Sólo hay que consultar el *smartphone*. Es algo asombroso.

Las últimas dos décadas, mi equipo del MIT Media Lab y yo hemos estudiado cómo nos afecta esto a nosotros y a nuestras interacciones, y cómo influyen estos efectos en los grandes cambios registrados por la sociedad. Por ejemplo, a través de la IA examinamos el patrón de interacción en comunidades, y nos preguntamos si los miembros de las comunidades hablan entre ellos, si hablan o no con el resto de la sociedad, si exploran o hablan con personas nuevas, o si hablan una y otra vez con las mismas personas.

Estudiando estos patrones hemos observado algo realmente sorprendente: se puede comprender la esencia de una comunidad, esas características humanas que antes eran imposibles de cuantificar, simplemente midiendo sus patrones de interacción y la IA. Los patrones de interacción nos cuentan si la comunidad está sana o enferma, si es pobre o rica, si es innovadora o si no avanza. Todo eso puede verse en el patrón de interacción entre las personas y con algoritmos de IA. Porque no es la tecnología la que dicta la esencia de la comunidad; son las personas y son las ideas que fluyen entre las personas lo que diferencia a una comunidad de otra.

Por otra parte, dándole la vuelta a este hallazgo científico podemos analizar el patrón de interacciones de esa comunidad a través de los coches o las bicicletas que circulan, a través del uso de las tarjetas de crédito o de los registros gubernamentales, y a partir de esos patrones podemos medir el lado humano de la comunidad. A partir de esas miguitas digitales y con IA se pueden identificar bolsas de pobreza, se pueden identificar barrios con un mayor riesgo de delincuencia, o barrios que se están convirtiendo en centros de innovación.

Por ejemplo, en Costa de Marfil, donde han sufrido recientemente una guerra civil, y donde no tienen censos, hemos visto que usando referencias de patrones de comunicación e IA se puede determinar qué poblaciones son pobres y cuáles son ricas. Y hacer este tipo de análisis sale barato. Eso significa tener un censo de la pobreza a tiempo real por sólo unos céntimos. Del mismo modo, se puede ver qué barrios tienen más posibilidades de convertirse en fuentes de innovación porque comunican mejor, y cuáles están en riesgo de registrar un aumento de la delincuencia porque cada vez están más aislados de las comunidades de los alrededores.

En realidad impresiona ver a través de la IA y la ciencia de los datos, que los patrones de comunicación son, aparentemente, el factor más importante a la hora de determinar las características humanas de una comunidad. Según parece no es el nivel de educación lo que más importa, ni tampoco el marco regulatorio: son las interacciones humanas. Antes de disponer de este tipo de referencias no entendíamos la importancia de los patrones de interacción humana. Había muchas teorías sobre por qué es diferente la delincuencia en distintos lugares, pero no tuvimos la respuesta hasta que empezamos a disponer de esas miguitas digitales que nos han permitido analizar los patrones reales de interacción.

Este cambio de «La mitad del mundo no ha hecho nunca una llamada telefónica» a «Podemos disponer de mapas de pobreza y de delincuencia a tiempo real» es lo que el secretario general de la ONU llama la Revolución de los Datos y la IA. Por primera vez en la historia del mundo tenemos realmente la posibilidad de ver lo que sucede en el mundo en forma de datos.

Durante toda la historia de la humanidad hemos sufrido epidemias que han diezmado la población, y los gobernantes, desde su capital, no se enteraban siquiera de lo ocurrido hasta después de que hubiera muerto la gente. En 1918, por ejemplo, hasta las sociedades más avanzadas hacían el seguimiento de la mortífera gripe registrando el número de cuerpos que llegaban al depósito. Pero ahora, gracias a una tecnología digital móvil que está por todas partes, a unos potentes ordenadores y a unas elaboradas técnicas estadísticas, podemos detectar las primeras señales de estos desastres e irlos siguiendo a medida que se desarrollan. Por ejemplo, la próxima vez que haya una epidemia de ébola, podremos usar montones de miguitas digitales y IA para hacer el seguimiento, frenarla y contenerla.

La Asamblea General de las Naciones Unidas ha decidido que los Institutos Nacionales de Estadística de cada país usen este tipo de métodos de análisis de

informaciones digitales de forma continuada por el bien de la sociedad. Eso significa que los 193 países miembros de la Asamblea General de la ONU deben comprometerse a que sus Institutos de Estadística comuniquen estas mediciones de «datos para el bien social» de modo que el mundo pueda tener más información sobre violencia, desigualdades, calidad de vida, propagación de enfermedades, etc., de forma más o menos continuada. Y así, por primera vez, podemos imaginarnos un mundo en el que haya de verdad transparencia y en que se responda del funcionamiento de las políticas de los gobiernos y de las intervenciones sociales, así como del alcance de la ayuda al desarrollo.

Estamos tan acostumbrados a que la situación real quede oculta tras una nube opaca que nos cuesta imaginar un mundo en el que por todas partes se hagan mediciones fiables y a tiempo real sobre la condición humana. Pero lo cierto es que estamos progresando en dirección a ese objetivo. Por supuesto, todos estos datos que podrían proporcionarnos recursos prácticos pero también transparencia, responsabilidad, un mejor gobierno y un mundo que tenga en cuenta a todos y cada uno de los seres humanos y se preocupe por ellos, también tienen un aspecto negativo.

El aspecto negativo son los delincuentes informáticos y las malas praxis comerciales, que no respetan las normas básicas de privacidad de datos, y a muchos les preocupa este lado oscuro del uso de los registros. Telefónica, por ejemplo, ha creado un índice con el que pregunta a personas de todo el mundo qué les parecen sus registros, y han observado que un 25 por ciento más o menos está de acuerdo con la ecología de datos actual. No obstante, otro 25 por ciento manifiesta serias dudas de que vayan a producir resultados positivos, y cree que sus informaciones, sumados al impacto de los algoritmos de la inteligencia artificial que los procesan, pueden provocar resultados negativos.

Así pues, ¿qué vamos a hacer con todos esos datos? Bueno, yo creo que tenemos que promover el uso de éstos con fines positivos, para comprendernos mejor a nosotros mismos, comprender nuestras comunidades y ayudar al mundo en general. Pero también tenemos que ser prudentes; tenemos que hacer los deberes. Pablo Rodríguez, autor de este libro, fue fundamental para que mi equipo del MIT, la Mozilla Foundation, el Open Data Institute del Reino Unido y Telefónica combinaran esfuerzos en la creación del Data Transparency Lab, para que los investigadores de todo el mundo puedan comprender mejor lo que se hace con nuestros registros, los algoritmos de IA y cómo hacer el mejor uso posible de ellos.

Este libro promueve esta actitud, el uso de los datos y la IA para el bien personal y para el bien de la sociedad, pero comprendiendo al mismo tiempo los peligros de un mundo rico en éstos, y cómo se puede construir un mundo que nos proteja de estos peligros. Los datos tienen poder. Los datos son valiosos. Este libro explora el potencial de éstos y de la inteligencia artificial, sus aplicaciones para el individuo y su potencial para transformar el mundo por completo, con una llamada a la prudencia

para evitar las trampas que se nos presentan, de modo que todos podamos disfrutar de un planeta mejor, con menos desigualdades y más oportunidades para todos.

PROF. SANDY PENTLAND
MIT Media Lab

Introducción

Corría el año 1997, en un laboratorio en el sótano del King's College de Londres. En el momento de entrar, estaba lloviendo y, siendo Londres, probablemente seguiría lloviendo en aquel momento. También habría oscurecido. Pero no había ventanas, así que no podía ver el exterior, y no había hablado con nadie desde mi llegada, salvo con el profesor Alan Rogers, supervisor de mi tesis. Llevaba trabajando sin hacer una pausa desde el momento en que había dejado mis cosas en la taquilla y cerrado la puerta del laboratorio. Estaba trabajando en un experimento relacionado con mi tesis, mediciones de alta resolución usando técnicas fotónicas. Parte de mi trabajo consistía en manipular lentes y láseres pero, sobre todo, implicaba estar sentado frente al terminal del ordenador jugando con algoritmos para contar fotones que volaban por ahí para poder señalar cambios de temperatura muy precisos que nos permitieran detectar tumores en fases muy tempranas.

Pero cada vez le encontraba menos encanto al trabajo de física y fotónica. En lugar de usar mi ordenador para procesar números, empecé a hacer pruebas con él, simplemente para ver cómo funcionaba y experimentar. Este ordenador había cambiado radicalmente mi modo de trabajar. Cuando estudiaba en España usábamos un antiguo lenguaje de programación llamado Fortran. Los datos se almacenaban en discos flexibles del aspecto de posavasos en los que sólo cabían migajas de ellos en comparación con lo que acumula hoy en día un reloj digital. Muchas veces los dejábamos trabajando toda la noche para que completaran los cálculos, y volvíamos al día siguiente como niños esperando los regalos del *tió* (El *tió* es una peculiar tradición catalana consistente en un tronco que es «alimentado» con golosinas por los niños durante el adviento. ¡En Navidad le dan golpes con palos mientras cantan una canción y el tronco «caga» regalos!).

En términos tecnológicos, los ordenadores que usé en España durante la carrera universitaria eran tan avanzados como troncos de madera en comparación con el ordenador con el que estaba jugando ahora en Londres. Y cada año los ordenadores se volvían más pequeños, más baratos, más potentes y más rápidos.

Yo también había ido cambiando a un ritmo similar. Hasta aquel momento me contentaba con trabajar solo en mis experimentos, dejando volar la imaginación a través de los algoritmos, y trasteando entre lentes y láseres con los dedos. Durante mi juventud en mi Asturias natal me sentía un bicho raro, un introvertido en tierra de extrovertidos. Los españoles, y los asturianos en particular, suelen ser tipos gregarios, que buscan el ruido y la compañía. Si entráis en cualquier bar de Oviedo os encontraréis gente charlando, haciendo esfuerzos por hacerse oír con el ruido del televisor de la esquina y la radio a todo volumen.

No es que fuera tímido. Es que me encantaba mi trabajo. Siempre me ha gustado disfrutar del placer de descubrir cosas nuevas. Pero el tipo de trabajo que estaba haciendo en Londres, pasando horas y horas solo haciendo cálculos matemáticos y experimentos, ya no me provocaba ninguna emoción. Había otra cosa que me llamaba más la atención: el ordenador e internet. Esa poderosa herramienta que podía hacer tantos cálculos, permitiéndome dejar volar la imaginación, en lugar de empantanarme en cálculos que ahora una máquina podía hacer por mí. Aquello suponía experimentar los efectos de tener una inteligencia artificial y aumentada a mi disposición.

Con el descubrimiento de los ordenadores y de internet cambió algo más. Descubrí gente muy interesantes. Extraño, ¿no? Normalmente tenemos esa imagen de los ingenieros informáticos trabajando solos por la noche, con el pálido rostro iluminado por el brillo de la pantalla. En aquella época, la de los albores de internet, no era así en absoluto. Era una comunidad de personas, en su mayoría científicos e ingenieros como yo (que ahora se llaman científicos informáticos), todos trabajando en una cosa completamente nueva, algo cuyo significado quizá sólo unos pocos de ellos entendían realmente. El objeto de su estudio era internet y la web, una red de comunicaciones, un almacén de conocimiento que acabaría volviéndose omnipresente en el mundo en una década.

Aquello creció tan rápidamente que, igual que un joven que de pronto crece quince centímetros, el estirón empezó a provocarle dolores. Estábamos estirando internet y llevándola al límite, y eso es en lo que yo trabajaba: en hacer que internet fuera más rápida y más eficiente para la cantidad de registros que transportaba, que aumentaba exponencialmente (desde complejas páginas web hasta vídeos en directo, o parches antivirus para evitar ciberataques).

Trabajé en algunos de los lugares más innovadores y pioneros de la investigación de internet, en Microsoft, en los Bell Laboratories y en el Instituto Federal Suizo de Tecnología. Descubrí y acabé formando parte de una comunidad de investigadores, científicos e ingenieros de datos y algoritmos que iban prácticamente improvisando, diseñando, construyendo y manteniendo algo nunca visto. Una comunidad que formaba parte de la historia de internet, que la transformó de un servicio lento y experimental a una maravilla a la velocidad de la luz. Fue una época emocionante; una época de muchísimo trabajo, de horarios infinitos, de grandes amistades y de innovaciones vertiginosas.

A los pocos años, ya hacía tiempo que había dejado el trabajo solitario del laboratorio londinense en el King's College, trabajando sobre oscuros problemas de la física teórica. Estaba creando y construyendo nuevos negocios. Creábamos y vendíamos compañías tecnológicas en Silicon Valley. Era profesor asociado en grandes universidades estadounidenses. Y aún no había cumplido los cuarenta años. Pero se acercaba otra transformación (y sin duda habrá más en el futuro). Y esa transformación es el objeto de este libro. Fue entonces cuando volví a España (¡y al

sol!) y (re)conecté con una comunidad, una comunidad globalizada de líderes en campos variados como la medicina, el liderazgo, la creatividad o la expendeduría.

Siempre he querido resolver problemas. Primero fueron teóricos; luego fueron tecnológicos. Ahora, con todo lo que he aprendido y con un extraordinario equipo, me dedico a problemas que pueden tener un tremendo impacto en la sociedad y que requieren grandes avances tecnológicos. ¿Qué pueden aportar los ingenieros informáticos como yo a problemas como la depresión? ¿O la pandemia de crisis global? ¿O la delincuencia? ¿O el liderazgo? ¿O el turismo?

He pasado gran parte de mi vida profesional estudiando y diseñando soluciones para enormes redes de telecomunicaciones. Redes que se extienden por continentes enteros y que sólo en los últimos cinco años han transportado más datos que todos los generados por la humanidad desde el inicio de la historia. Pero me he dado cuenta, a través de conversaciones, reuniones casuales y a veces por simples ideas que se me han ocurrido, que hoy en día el conocimiento y la experiencia que han ganado los ingenieros informáticos durante los vertiginosos años de aparición de internet y con los recientes avances en *big data* e inteligencia artificial (IA) pueden aplicarse a otras disciplinas y al nivel personal, con resultados extraordinarios. Es como enrolar a un constructor de barcos en un proyecto de diseño de una catedral (algo que en realidad ya se está haciendo, en los trabajos de construcción de la Sagrada Familia de Barcelona). No solo supone la incorporación de otro punto de vista, sino también de herramientas específicas y procesos desarrollados para un campo pero adaptables a otro.

Y esta influencia de diferentes tecnologías, los datos y la inteligencia artificial en la sociedad están creando grandes cambios. La mayoría del dinero que has gastado en los últimos veinticinco años en libros, ocio, educación, información salud y transporte se está desmaterializando gracias a la tecnología, los ordenadores e internet, que hacen que tu dinero llegue más lejos para que puedas usarlo con otros fines. La tecnología está haciendo posible que todas estas necesidades cuesten mucho menos, probablemente lo que cuesta una conexión a internet. Muy pronto contarás con los mejores laboratorios de análisis médicos en el teléfono, a través de sensores que detectarán y analizarán tus signos vitales o te analizarán el ADN; no será necesario trasladarse tanto para acceder a los mejores médicos y especialistas; ya tienes todas las enciclopedias posibles en la punta de los dedos, acceso a las noticias, información y ocio, y los sistemas de computación más potentes a tu alcance. Recuperarás tiempo libre para hacer mejor lo que más te guste y lo que mejor se te dé, ganando así tiempo para hacer más actividades cada vez y mejorar exponencialmente.

Internet ha hecho que dispongamos de datos en abundancia, que estén por todas partes y que sean mucho más valiosos. Tanto si sales a correr como si ves la tele o simplemente estás sentado en el coche en un atasco, prácticamente cada actividad que desarrollas crea un rastro digital, más materia prima para las destilerías de registros.

Ahora que dispositivos tan diversos como relojes o coches conectan con internet, el volumen va en aumento: hay quien calcula que un coche que se conduzca solo generará 100 gigabytes por segundo.

Por su parte, las técnicas de inteligencia artificial sacan un valor añadido a estas informaciones. Los algoritmos pueden predecir cuándo un cliente está listo para comprar, cuándo hay que revisar un motor a reacción o cuándo una persona corre peligro de desarrollar una enfermedad. Hoy en día, gigantes industriales como General Electric o Siemens se venden como empresas de datos.

Y esto es importante porque tus datos personales contienen tu alma. Éstos son un espejo de ti mismo, de tus emociones. La mayoría piensa que los datos son fríos, impersonales, pero pueden revelar tu lado emocional, pueden ayudarte a crecer. Los datos y la inteligencia artificial se están convirtiendo en el nuevo psicólogo, en el nuevo *coach*; ellos nos ayudan a prevenir nuestras enfermedades, a reaccionar durante desastres naturales y a resolver los problemas sociales más acuciantes.

Este libro cuenta la historia de las personas que he conocido y de las colaboraciones en que he participado para resolver problemas de muy diversos campos, utilizando registros y la inteligencia artificial. Una nueva innovación que crea impacto a nivel global y es sostenible y se preocupa de la conciencia humana. Ahora aquel sótano de Londres queda muy lejos. Visito hospitales, organizaciones internacionales de salud, grandes equipos de liderazgo y creatividad, viajo a las profundidades de los centros de datos y al centro de la red y voy aprendiendo sobre salud, alimentación, energía, sobre economía mundial o sobre lo que hace feliz a la gente. Aplico todo lo que he aprendido, y sigo aprendiendo, para ayudar a resolver los problemas de médicos y enfermos, ingenieros de redes, activistas en defensa de los derechos humanos, organizaciones humanitarias, entrenadores de fútbol, artistas y chefs de cocina, empresas sociales, organizaciones sin ánimo de lucro, o tus propios problemas, para hacer del mundo un lugar mejor. Y todo ello está sucediendo con la revolución del «alma de tus datos», y de la inteligencia artificial.

DOCTOR PABLO RODRÍGUEZ

1

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 condiciones 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.

Los inicios de la inteligencia empresarial

1958 Hans Peter Luhn, investigador de IBM, define la inteligencia empresarial como «la capacidad de percibir las interrelaciones de los hechos presentados de modo que puedan guiar la actuación [de la empresa] hacia una meta deseada».

1962 Se dan los primeros pasos hacia el reconocimiento de voz, cuando William C. Dersch, ingeniero de la IBM, presenta la Shoebox Machine en la Feria Mundial de 1962. La Shoebox puede interpretar números y dieciséis palabras pronunciadas en inglés y convertirlas en información digital.

1964 Un artículo publicado en *New Statesman* hace referencia a la dificultad para gestionar la cada vez mayor cantidad de información de que disponemos.

La aparición de los grandes centros de datos

1965 El gobierno de Estados Unidos planea la creación del primer centro de datos del mundo para registrar 742 millones de devoluciones de impuestos y 175 millones de juegos de huellas digitales en cinta magnética.

1970 Edgar F. Codd, matemático de IBM, presenta su propuesta para una «base de datos relacional». El modelo proporciona el marco operativo que usan muchos servicios de registros modernos hoy en día, para archivar información en un formato jerárquico de modo que pueda acceder cualquiera que sepa lo que busca. Hasta este momento, para acceder a los datos de los bancos de memoria de un ordenador hacía falta la participación de un experto.

1989 Se usa, posiblemente por primera vez, el término macrodatos (*big data*) con el sentido que se le da hoy. Erik Larson, escritor superventas internacional, escribe un artículo en Harper's Magazine en el que especula sobre el origen del correo basura que recibe.

La aparición de internet

1991 El científico informático Tim Berners-Lee anuncia el nacimiento de lo que se convertiría en la Web tal como la conocemos hoy. En un post colgado en el grupo de Usenet alt.hypertext indica las especificaciones para una red mundial de datos interconectada, accesible para cualquier persona desde cualquier lugar.

1996 En su libro *The Evolution of Storage Systems*, de 2003, R.J.T. Morris y B.J. Truskowski afirman que este es el año en que el almacenaje digital empezó a ser más rentable que el almacenaje en papel.

1997 Michael Lesk publica su artículo «How Much Information is there in the World?» («¿Cuánta información hay en el mundo?»), teorizando que «quizá no sea exagerado suponer» una cifra de 12.000 petabytes. También señala que incluso en esta fase temprana de su desarrollo, la web va multiplicando su tamaño por diez cada año, y demuestra que gran parte de los datos nunca los verá nadie, por lo que no aportan ninguna información útil.

Google Search también hace su debut este año, y durante al menos dos décadas será sinónimo de búsqueda de registros en la red.

Las primeras ideas sobre macrodatos

1999 Apenas un par de años después, el término *big data* aparece en el artículo «Visually Exploring Gigabyte Datasets in Real Time» («Exploración visual de conjuntos de datos de gigabytes en tiempo real»), publicado por la Association for Computing Machinery. Una vez más, se lamenta la propensión a almacenar una cantidad significativa de referencias sin que haya posibilidad de analizarla adecuadamente. El artículo cita a Richard W. Hamming, pionero de la informática, que había dicho «El objetivo de la computación es la comprensión, no los números».

Posiblemente también sea el momento en que se usa por primera vez el término «Internet de las cosas» para describir el creciente número de mecanismos en línea y el potencial que tienen para comunicar unos con otros, en muchos casos sin un «intermediario» humano. El término se usa como título de una presentación realizada en Procter and Gamble por Kevin Ashton, pionero de la identificación por radiofrecuencia.

- 2000** Con el estudio «How Much Information?» («¿Cuánta información?»), Peter Lyman y Hal Varian intentan cuantificar por primera vez la cantidad de información digital del mundo, y su ritmo de crecimiento. Concluyen: «La producción total anual de contenidos en formato impreso, filmado, óptico y magnético del mundo requeriría aproximadamente 1.500 millones de gigabytes para su almacenaje. Eso equivale a 250 megabytes por persona, es decir por cada hombre, mujer y niño de la Tierra».
- 2001** En su estudio «3D Data Management Controlling Data Volume, Velocity, and Variety» («Gestión de datos en 3D, el control del volumen, la velocidad y la variedad de datos»), Doug Laney, analista de la Gartner Inc., define tres de las que acabarán siendo las características más comúnmente aceptadas de los macrodatos.

La web 2.0 aumenta los volúmenes de datos

- 2005** Los comentaristas anuncian que estamos presenciando el nacimiento de la «Web 2.0», la red generada por el usuario en la que la mayoría del contenido lo generarán los usuarios de los servicios, y no los proveedores de los mismos.

También es el año en que se crea Hadoop, la plataforma de código abierto creada específicamente para el almacenaje y el análisis de conjuntos de macrodatos. Su flexibilidad lo hace especialmente útil para gestionar registros desestructurados (voz, vídeo, texto sin formato, etc.) que generamos y recopilamos cada vez en mayor cantidad.

Se adopta el uso actual del término «big data» (macrodatos)

- 2007** La revista *Wired* acerca el concepto de los metadatos a las masas con su artículo «The End of Theory The Data Deluge Makes the Scientific Model Obsolete».
- 2008** Los servidores del mundo procesan 9,57 zettabytes (9,57 billones de gigabytes) de información, el equivalente a 12 gigabytes de información por persona y día), según el informe «How Much Information?» de 2010. En el artículo «International Production and Dissemination of Information», se calcula que este año se producirán 14,7 exabytes de nuevos datos.
- 2009** La empresa media con más de 1.000 empleados de Estados Unidos almacena más de 200 terabytes de informaciones, según el informe «Big Data The Next Frontier for Innovation, Competition, and Productivity» del McKinsey Global Institute.
- 2010** Eric Schmidt, presidente ejecutivo de Google, da un discurso en una conferencia en la que dice a los participantes que cada dos días creamos tantos datos como los que se han recogido desde el inicio de la civilización humana hasta el año 2003.

2011 El informe McKinsey advierte de que en 2018 Estados Unidos se enfrentará a una carencia de entre 140.000 y 190.000 profesionales expertos en datos y de que habrá que resolver problemas como la privacidad, la seguridad y la propiedad intelectual de ellos antes de que seamos conscientes del valor real de los macrodatos.

2014 De los ejecutivos que trabajan con Accenture, el 88 por ciento de los encuestados por GE afirman que el análisis de macrodatos está entre las prioridades de su empresa.

Lo que nos enseña Bernard es que los macrodatos no son un fenómeno nuevo o aislado, sino algo que forma parte de nuestra larga evolución de captura y uso de datos. Al igual que otros avances clave en el almacenamiento estos últimos, en los ordenadores y en internet, los macrodatos son un paso más que aportará cambios al modo en que gestionamos nuestras empresas y la red. Al mismo tiempo, sentará los cimientos sobre los que se forjarán muchos avances.

El McKinsey Global Institute (MGI) estudió los macrodatos en cinco dominios: la asistencia sanitaria en Estados Unidos, el sector público en Europa, las ventas al público en Estados Unidos y la manufactura y los datos de localización personal en todo el mundo. Los macrodatos pueden generar valor en todos ellos. Por ejemplo, un vendedor que use los macrodatos a fondo podría aumentar su margen de beneficios en más de un 60 por ciento.

Controlar grandes cantidades de información también puede suponer una gran ventaja en el sector público. Si el sistema de asistencia sanitaria de Estados Unidos usara los macrodatos de forma creativa y efectiva para potenciar la eficiencia y la calidad, el sector podría crear más de 300.000 millones de dólares en valor añadido cada año. Dos tercios de esa cantidad sería en la reducción del gasto sanitario nacional en un 8 por ciento aproximadamente. En las economías desarrolladas de Europa, los gestores gubernamentales podrían ahorrar más de 100.000 millones de euros en mejoras de la eficiencia operativa sólo con el uso de macrodatos, sin incluir el uso de las referencias para la reducción del fraude y de los errores y la mejora en la recaudación de impuestos.

Para que los datos sean útiles y se entiendan, deben ser recopilados y transportados a través de potentes redes de todo el mundo y, afortunadamente, también en ese campo hemos hecho grandes progresos. Por otra parte, éstos por sí solos no significan nada. Suele ser necesario establecer complejas relaciones entre ellos usando redes y conexiones que hay que analizar a fondo (por ejemplo, nuestras relaciones sociales en línea, los «me gusta» de nuestras páginas y de las páginas a las que llevan las páginas que nos gustan a nosotros, el patrón de actividad de telefonía móvil en una ciudad y cómo se mueve la multitud durante un evento, la red creada por los tuiteos al retuitearlos, etc.). Cuando observamos nuestros datos personales con

una visión de grupo, y analizamos cómo utilizarlos en la red, nos encontramos con cosas muy sorprendentes.

Si se utilizan los registros introducidos en la red, los «me gusta» de Facebook y los vínculos que has clicado, tu lista de reproducción de Spotify puede resultar muy útil. En lugar de tener que preguntarte si te gusta la poesía, ahora se puede echar un vistazo a tus «me gusta» de Facebook y obtener exactamente la misma información. Los datos y algoritmos que analizan cientos de miles de éstos pueden tener en cuenta pequeñas informaciones, sus interconexiones y los patrones de conexión en la red, combinarlos y predecir cosas muy complicadas. Si los datos están conectados, se pueden convertir datos aparentemente no relacionados en grandes historias sobre ti mismo y sobre tu potencial futuro.

No tiene nada de sorprendente echar un vistazo a tu lista de Spotify y ver que si tienes un montón de álbumes de Lady Gaga, quizá te guste ella y su tipo de música. Lo que resulta revolucionario del nuevo mundo de los datos y los algoritmos es que pueden usar tus preferencias musicales, o tus preferencias literarias, sus patrones de conexión en la red, y de esas relaciones aparentemente inocentes pueden hacer predicciones muy exactas de tu religiosidad, de tu rendimiento económico futuro, de los riesgos de tu empresa, de tu potencial de liderazgo, de tus opiniones políticas, etc. Esos modelos de relación de los registros en la red pueden generar mecanismos muy potentes para comprender el futuro, a las personas y la sociedad, y ofrecen unas oportunidades nunca vistas.

Google, el motor de búsqueda más potente del mundo, basa toda su inteligencia en uno de esos algoritmos de referencias en red, el PageRank, también conocido como el algoritmo algebraico lineal de los 25.000 millones de dólares. Este algoritmo es capaz de predecir cuáles son las páginas más importantes y cuáles deberían aparecer en primer lugar en el resultado de tus búsquedas, basándose en la red de datos creada a partir del modo en que están relacionadas unas páginas con otras, de qué sitios web dirigen a qué otros, y de lo importante que sean, lo cual se decide según quiénes sean los que los dirigen a ellos. El PageRank funciona contando el número y la calidad de los vínculos que llevan a una página para hacer un cálculo aproximado de la importancia de ese sitio web en relación con todo el conjunto de páginas de la red. Se supone que los sitios web más importantes probablemente tendrán más vínculos procedentes de otros sitios web. Las redes de datos y su estructura están en el centro de la revolución de esto últimos, el aprendizaje a partir de ellos y la inteligencia artificial.

El poder de las redes de datos

Barcelona, la ciudad donde vivo, no tiene puentes que crucen la ciudad. A diferencia de Königsbergin, población monástica prusiana que, tras quedar

prácticamente destruida durante la segunda guerra mundial y ocupada posteriormente por la Unión Soviética, fue rebautizada con el nombre de Kaliningrado. Fue fundada en 1255 por la Orden Teutónica y fue capital de la Prusia Oriental hasta 1945. Durante siglos fue un centro de aprendizaje donde vivieron Goldbach, Hilbert, Kant, y Wagner.

El río Pregel atravesaba la ciudad. Separaba dos grandes zonas en la costa y dos grandes islas. Había siete puentes que conectaban las diversas zonas, que dieron pie a un conocido problema matemático: los residentes de Königsberg se preguntaban si era posible recorrer toda la ciudad cruzando cada uno de los siete puentes una vez y no más.

Leonhard Euler (1707-1783), nacido en Basilea (Suiza), estaba destinado a convertirse en pastor religioso, como su padre. Estudió hebreo y teología en la universidad, pero en privado recibió clases de matemáticas. En 1727 consiguió un trabajo en el departamento de medicina de la Universidad de San Petersburgo, pero en el caos resultante tras la muerte de la emperatriz Catalina I, consiguió colarse en el departamento de matemáticas. En 1733 Euler estaba casado y tenía trece hijos, cinco de los cuales consiguieron sobrevivir hasta la edad adulta. En 1741 se trasladó a Berlín, donde vivió veinticinco años.

En toda su vida publicó más de 500 libros y artículos, y otros 400 se publicaron de forma póstuma. Cabe destacar que definió, entre otras, las notaciones i , p , e , \sin y \cos . Perdió la vista en ambos ojos, pero eso no hizo más que aumentar su productividad. «Ahora tengo menos distracciones», dijo, ¡y en 1736 hizo historia en las matemáticas al hablar de redes! (como internet, la web, las redes sociales, etc.).

Y lo hizo encontrando una solución al problema de los puentes de Königsberg (es decir, el de cómo pasar por todos los puentes sin pasar más de una vez por cada uno de ellos). Se dio cuenta de que no importaba cómo caminabas por el terreno, la forma de las islas, la longitud o la anchura de los puentes, ni dónde estaban los puentes exactamente. Sólo importaba cuántos puentes había entre cada pedazo de tierra, y en qué orden los cruzabas. Lo único que importaba era la estructura de red creada por los puentes. ¡Y fue un resultado espectacular!

La ciudad prusiana de Königsberg (ahora Kaliningrado, en Rusia) se encontraba a ambos lados del río Pregel y contaba con dos grandes islas, interconectadas entre sí y con la costa por siete puentes. El problema era idear un recorrido por la ciudad que supusiera cruzar cada puente una vez y sólo una, con estas condiciones: que a las islas sólo se podía llegar por los puentes, y que cada puente que se pisara se recorriera hasta el final. No era necesario finalizar el recorrido en el punto de partida.

Tras muchos largos paseos intentando encontrar la solución y mucho ensayo y error, Euler decidió idear una demostración matemática que dejara claro que el problema de Königsberg no tenía solución; nadie podía atravesar los puentes una sola vez y volver al punto de partida.

Para hacerlo, trazó un esquema de la red de Königsberg que contenía únicamente la información que necesitaba —un punto para cada una de las masas de tierra separadas por los ríos— y conectó los puntos con una línea en el caso en que hubiera un puente que unía las masas de tierra. ¡Lo que dibujó ahora lo denominamos una red, y Euler fue el primero en hacerlo! Demostró que cruzar todos los puentes de Königsberg una sola vez era imposible (a diferencia, por ejemplo, de los puentes que cruzan la isla de Notre Dame, en París).

En la historia de las matemáticas, la solución de Euler al problema de los puentes de Königsberg es considerada el primer teorema de la teoría de redes, tema considerado hoy básico para comprender los datos de la web e internet. Porque igual que internet o la web, que es una red de ordenadores y registros, una ciudad es una red de calles, avenidas y líneas de metro.

Euler generó un algoritmo para encontrar una solución que permitiera cruzar todos los puentes sólo una vez, y su algoritmo ha tenido muchas aplicaciones a lo largo del tiempo y ha ayudado a mejorar ciudades de todo el mundo. Por ejemplo, los ayuntamientos de las ciudades gastan millones de dólares cada año en servicios como el barrido de las calles o la retirada de nieve. Las soluciones como la ideada por Euler pueden ayudar a las ciudades a ahorrar dinero o a idear algoritmos eficientes para mantener limpias las calles.

La teoría de redes que han encontrado numerosas aplicaciones en las ciudades, el transporte y el almacenaje, la planificación y organización de agendas, el diseño de internet, los algoritmos de búsqueda en la web, la inteligencia artificial, y el análisis de los sistemas eléctricos.

La solución de Euler también puede ser utilizada por las empresas de correo que quieren trazar una ruta que no suponga pasar dos veces por el mismo sitio, o por quienes gestionan el trabajo de los jardineros que se ocupan de los parques públicos. Pongamos que hay numerosos senderos en el parque, y que cada sendero tiene una hilera de flores al lado. El jardinero tiene que regar todas las flores una vez únicamente, y tiene que hacerlo en el mínimo tiempo posible. La solución puede estar en la ciencia de las redes y de los datos. Otro ejemplo de aplicación podría ser el pintado de rayas en una nueva calle. Pongamos que hay una ciudad en la que todas las calles son nuevas y que aún hay que pintarlas. Los pintores necesitan pintar una raya en cada calle, y sólo una. Así que tienen que encontrar el modo de poder recorrer cada vía una sola vez (no quieren pintar más que lo absolutamente necesario).

Dos de los siete puentes originales no sobrevivieron a los bombardeos de Königsberg durante la segunda guerra mundial. Otros dos fueron demolidos posteriormente y sustituidos por una moderna carretera. Los otros tres puentes siguen ahí, aunque sólo dos de ellos son de la época de Euler (uno fue reconstruido en 1935). Así, en el año 2000 había cinco puentes en Kaliningrado, y actualmente, con la nueva disposición de los puentes en la red urbana, sí es posible encontrar una solución al antiguo problema. No obstante, lo más importante es que este caso abrió la

posibilidad de comprender otras redes de macrodatos como la web, internet, Facebook, LinkedIn o Twitter. Y eso, desde luego, es fascinante.

Podemos encontrar el origen matemático de las redes en los primeros trabajos de Leonard Euler sobre teoría de grafos, en el siglo XVIII. Desde entonces, las redes se han ido convirtiendo en un concepto interdisciplinar que se apoya en disciplinas como la física, la matemática, la informática, la biología y la economía.

Pero hoy en día lo más importante para nosotros es el auge de las redes de datos en las últimas dos décadas. Internet y la web, con su espectacular crecimiento e impacto, son las redes que han despertado más la imaginación de numerosos expertos en redes. Ahora que disponemos de una capacidad mucho mayor para analizar esos datos, podemos estudiar el potencial de las interacciones en línea en una escala masiva, global, con la esperanza de hacer nuevos descubrimientos posiblemente de un valor enorme tanto para los estudiosos de las redes como para el mundo en general.

Por ejemplo, podemos estudiar cómo interactúan las comunidades (qué miembros hablan entre sí, si se relacionan con otras comunidades, con qué frecuencia, etc.). Estudiando esos patrones, los científicos del MIT Media Lab han descubierto algo realmente sorprendente: se puede comprender la esencia de una comunidad, esas características humanas que hasta ahora resultaban imposibles de calcular, simplemente midiendo su patrón de interacciones. Los patrones de interacciones nos dicen si la comunidad está sana o enferma, si es rica o pobre, si es creativa o si no muestra actividad. Todo eso puede verse en el patrón de interacción entre las personas. Son las personas y las ideas que fluyen entre las personas las que diferencian una comunidad de otra.

Si tomamos este hallazgo científico y lo ponemos en diferentes entornos, nos damos cuenta de que podemos analizar el alma de los registros de los individuos y de una comunidad, el rastro digital de sus interacciones, y podemos identificar bolsas de pobreza, de desempleo, podemos identificar barrios más susceptibles de sufrir más inseguridad, y barrios que se están convirtiendo en motores de riqueza.

Del mismo modo, las redes sociales como Facebook, Twitter o LinkedIn son redes de datos que generan enormes cantidades de registros y tienen propiedades «mágicas» que conectan a las personas de todo el mundo a través de su red de contactos sociales, aunque ningún humano conozca la estructura de la red. Ésta es una propiedad única que hace que los datos fluyan por todo el mundo a una velocidad increíble. Y crea historias que nos unen a unos con otros para conseguir grandes cosas.

Interconectando Datos: propiedades mágicas

Hay pruebas del poder de las redes de datos por todas partes. Como humanos, tenemos la habilidad de compartir historias y cooperar de un modo efectivo con extraños, creando redes y comunidades que nos han convertido en la especie más avanzada del planeta. Hemos construido redes de interacción muy sofisticadas para librar guerras, luchar contra las hambrunas y contra las enfermedades. El poder de verdad siempre ha estado en las redes. Los seres humanos, por sí solos, no son capaces de hacer casi nada. Es lo que pueden hacer en grupo lo que les ha permitido hacerse con el control del planeta. Hoy en día estas agrupaciones —corporaciones, religiones, estados— son parte de una amplia red de flujos de información y registros interconectados. Queremos creer que nuestras vidas tienen algún significado objetivo, y que nuestros sacrificios tienen algún valor, más allá de las historias que tenemos en la cabeza. Sin embargo, la verdad es que la vida de muchas personas sólo tiene sentido en el entorno de la red de historias que se cuentan unos a otros. Se crea significado cuando muchas personas tejen juntas una red de historias comunes, de conocimientos y registros adquiridos y de interacciones.

Para poner a prueba el poder de estas redes y nuestra capacidad de cooperar y construir sistemas complejos, en 1967 Stanley Milgram decidió hacer un experimento que pondría del revés el mundo de las redes de datos y que propiciaría grandes progresos. Milgram eligió gente al azar de ciudades de Estados Unidos para trazar redes sociales y ver si existía alguna conexión mágica entre ellas. Si así era, se podría acelerar tremendamente el proceso de analizar todos esos patrones de redes de datos y esas interacciones.

En el experimento «Six degrees of separation» (Seis grados de separación), Milgram envió varios paquetes a 160 personas escogidas al azar que vivían en Omaha (Nebraska), pidiéndoles que enviaran el paquete a un amigo o conocido que pensarán que pudiera hacer llegar el paquete más cerca del destinatario final, un corredor de bolsa de Boston.

Cada uno de los «iniciadores» recibió instrucciones de enviar una carpeta a un receptor a través del servicio de correos, pero siguiendo algunas reglas. Los iniciadores sólo podían enviar el sobre a alguien que conocieran personalmente y con quien se tutearan. Con el envío, debían dar instrucciones al receptor para que reenviara el sobre a un conocido dándole las mismas instrucciones, con la esperanza de que esta persona conociera por casualidad al destinatario final.

Dado que los iniciadores sólo conocían el nombre y la dirección del destinatario final, la labor parecía imposible. Milgram hizo un seguimiento del progreso de cada cadena a través de un sistema de acuse de recibo de los envíos, que le permitía seguir el rastro a cada carta. Sorprendentemente, se encontró con que la primera carpeta llegaba a su objetivo en sólo cuatro días, y pasando únicamente por dos conocidos intermedios. En total, Milgram observó que la distancia entre el primer remitente y el destinatario final se cubría de media con cinco conocidos intermedios.

De hecho, esta propiedad mágica de determinados tipos de redes hace posible que internet envíe una cantidad enorme de datos y que dirija el tráfico eficientemente sin necesidad de recurrir a un sistema de coordinación central. Lo mismo es aplicable a los algoritmos usados por los motores de búsqueda para examinar rápidamente cientos de millones de páginas y dar buenos resultados de búsqueda en un instante. De otro modo sería prácticamente imposible. También hace posible que comunidades de personas alejadas interactúen de forma eficiente y puedan realizar tareas complejas y conseguir objetivos (por ejemplo, cooperar para librar guerras, combatir enfermedades o lanzar una iniciativa empresarial). Esta propiedad descubierta por Milgram de construir complejas redes posibilita que grupos de personas muy distantes entre sí compartan sueños y aficciones de un modo muy efectivo, y que puedan cooperar en pos de un objetivo común. La naturaleza de algún modo también «dirige los datos» de un modo eficiente sin necesidad de un sistema de control central. Lo mismo ocurre en nuestro cerebro, que es un aleccionador ejemplo de cómo transmitir efectivamente señales de datos específicos a determinados puntos del cuerpo, sin que ninguna neurona tenga una visión general de la red de conectividad interneuronal en su conjunto.

Con estos y otros avances, vemos que internet, y sus algoritmos de registros, adquieren una capacidad aparentemente infinita, yendo más allá de los límites actuales de la ciencia, de la informática y de la ingeniería para crear una red de conexiones de informaciones que pueden trabajar a velocidades próximas a la de la luz, y con una inteligencia propia.

También asistimos a un auge de las máquinas y mecanismos conectados a internet, que enviarán y recibirán cantidades masivas de referencias, al tiempo que transforman el panorama de la industria con una amplia gama de impresoras 3D, revertiendo la actual tendencia de los países desarrollados de derivar la fabricación a países donde resulta más barato.

Nos encontramos en un punto en que tenemos una visión futurista de los datos y de la información en red que podría sonar a ciencia ficción, más que a análisis y pensamiento tecnológico racional; no obstante, es un momento único, en él confluyen elementos tecnológicos comunes y necesidades humanas (de consumo y de empresa) que impulsan unas oportunidades de crecimiento exponencial.

Científicos de datos: los gurús de los macrodatos

Hoy en día el análisis de estas complejas redes de interacciones entre personas, ordenadores, empresas, comunidades en línea, ciudades, etc., los realizan los científicos e ingenieros expertos en macrodatos e inteligencia artificial, a los que se les requiere que extraigan patrones y respuestas a partir de grandes conjuntos de registros que antes no podían ser analizados. En los macrodatos y la inteligencia

artificial han influido la aparición de una capacidad de almacenamiento masivo en la nube, de sofisticados algoritmos que permiten extraer información de aquellos mismos, y de la disponibilidad de grandes volúmenes de datos digitales desde hace una década más o menos. Ese es el motivo por el que se ha dado tanta importancia a la posibilidad de diseccionar esos datos en busca de correlaciones que puedan darnos información y nuevas perspectivas.

Hoy en día, el estudio de los datos y la inteligencia artificial está considerada la profesión más «sexy» del siglo XXI. ¡Nada menos! Básicamente, lo que hacen los profesionales de los datos es hacer descubrimientos mientras están sumergidos en ellos. Es su método preferido de moverse por el mundo que les rodea. Se encuentran a gusto en el reino digital, y son capaces de dar una estructura a grandes cantidades de referencias e informes, posibilitando su análisis. Identifican ricas fuentes de datos, las combinan con otras fuentes de éstos potencialmente incompletas y limpian el conjunto resultante. En un panorama competitivo en que no dejan de aparecer nuevos desafíos y en que no dejan de aparecer datos, los expertos en ellos ayudan a los responsables de tomar decisiones a pasar de un análisis *ad hoc* a un diálogo sostenido con los datos.

Los expertos en datos se dan cuenta de que se enfrentan a limitaciones técnicas, pero no permiten que eso limite su búsqueda de soluciones. A medida que hacen descubrimientos, comunican lo que aprenden y comunican lo que ello implica para aplicarlo en nuevas iniciativas empresariales. En muchos casos son creativos en la exposición de la información de forma visual y consiguen que esos patrones sean claros e interesantes. Asesoran a los ejecutivos y a los directores de producción sobre las consecuencias que tienen esos registros en los productos, procesos y decisiones.

Dado que el suyo es un campo novedoso, a menudo les toca a los técnicos en datos crearse sus propias herramientas e incluso llevar a cabo investigaciones al estilo académico. Hoy en día sus servicios son solicitados por muchas empresas, como Telefónica, Google, Amazon, Microsoft, Walmart, eBay, LinkedIn o Twitter.

¿Qué tipo de persona es la que hace este trabajo? ¿Qué habilidades hacen que un técnico de datos tenga éxito? Podemos imaginárnoslo como una mezcla de pirata informático, analista, comunicador y asesor de confianza. Es una combinación tan increíblemente potente como poco habitual.

La habilidad universal y más básica de los técnicos de datos es el dominio de la codificación y los algoritmos, pero lo más importante es su capacidad para comunicar con un lenguaje que puedan entender todos los que invierten en ellos, y demostrar la extraordinaria habilidad que supone contar historias con referencias, sea verbalmente, visualmente o —en el mejor de los casos— de las dos formas. Pero podríamos decir que el rasgo dominante entre los técnicos de datos es una gran curiosidad, el deseo de explorar bajo la superficie de un problema, encontrar las cuestiones que dan origen al mismo y destilarlas en una serie de hipótesis muy claras que puedan ponerse a prueba.

Esto a menudo conlleva un pensamiento asociativo, característica de los científicos más creativos de todos los campos. Por ejemplo, tenemos constancia de un experto en datos e IA que, estudiando un problema de fraudes, se dio cuenta de que había una analogía con el problema de secuenciación del ADN. Uniendo esos dos mundos tan dispares, él y su equipo consiguieron idear una solución que redujo las pérdidas por fraude espectacularmente.

Hal Varian, director económico de Google, dijo una vez: «El trabajo más sexy de los próximos diez años será el análisis de registros y la inteligencia artificial. La gente cree que estoy de broma pero, ¿quién habría dicho que el trabajo más sexy de los años noventa iba a ser el de ingeniero informático?».

Si «sexy» significa tener cualidades únicas muy demandadas, los expertos en datos ya cumplen la condición. Son tipos desafiantes y salen caros, por lo competitivo que es el mercado que requiere sus servicios. Y resulta difícil retenerlos, básicamente porque no hay mucha gente que combine como ellos la formación científica con la técnica informática y analítica.

Los científicos de datos de hoy en día recuerdan a los analistas cuantitativos de los años 1980 y 1990. Por aquel entonces, el sector de la banca y los fondos de cobertura empezó a integrar a personas con formación en física y matemáticas para que idearan algoritmos y estrategias de datos completamente nuevos. A partir de entonces muchas universidades empezaron a ofrecer másters en ingeniería financiera, lo que dio origen a una segunda generación de talentos, más accesibles para las grandes empresas.

La evolución registrada en el mundo de los macrodatos y la IA no muestra indicios de que vaya a frenarse. Las empresas que, por falta de talento, se queden fuera de esta tendencia en su fase incipiente se arriesgan a quedar por detrás de sus competidores. Hay que pensar en los macrodatos y la IA como en una gran ola que se está generando ahora mismo, y que empieza a crecer. Si se quiere aprovechar la ola, se necesita contar con gente que sepa montarse en ella.

Los macrodatos y «la nube» no son la misma cosa, aunque de algún modo estén relacionados. Los macrodatos implican el análisis de grandes volúmenes de referencias diversos prácticamente en tiempo real (volumen, velocidad, variedad). Y casi todo esto los técnicos de datos lo hacen en la nube.

Empecemos con los macrodatos, o *big data*. No es sólo la descripción de los datos cuando se vuelven «*big*» (pongamos terabytes al día, por ejemplo). También se trata de cómo organizar éstos, de cómo clasificarlos, de qué tecnologías usamos para almacenarlos y recuperarlos, y de muchas otras facetas. Los macrodatos son un concepto más amplio, no sólo datos que simplemente son grandes en tamaño. Suponen, más bien, un nuevo paradigma de recopilación, almacenamiento, manipulación y extracción de significado de registros de diferentes tipos.

En cuanto a la computación en la nube, es el paradigma de ofrecer servicios de computación sobre la marcha. Todo (o casi) queda desmaterializado, y no hay que preocuparse demasiado por la configuración antes de empezar.

Tampoco hay necesidad de preocuparse por la gestión de una torre de servidores antes de empezar a usar herramientas de gran potencia. Por ejemplo, Dropbox te permite almacenar documentos y acceder a ellos mediante una simple conexión a internet. Eso es un servicio en la nube.

También puedes usar un servidor de alquiler de Amazon (AWS) o de Google (Compute Engine) para realizar tareas de gran coste computacional (invertir en una matriz de $10^{15} \times 10^{15}$, pongamos).

Antes de la aparición de la nube, estas tareas salían caras, suponían un desafío técnico y sólo estaban al alcance de unos pocos. Ahora se han democratizado y trabajar con macrodatos es algo asequible para cualquiera a través de la nube. Los *big data* representan el contenido, y la computación en la nube son su infraestructura.

Estos dos conceptos están interconectados. De hecho, es gracias a la computación en la nube que los técnicos de datos pueden procesar los macrodatos de tu ordenador. Y es la combinación de ambos lo que ha hecho que los macrodatos sean un concepto tan importante y el motivo por el que hoy en día es mucho más fácil que nunca lanzar una empresa tecnológica.

Desde su nacimiento, la tecnología de la información sólo ha estado al alcance de las empresas tecnológicas, las grandes organizaciones, los gobiernos y las instituciones académicas. Así fue hasta la aparición de la computación en la nube, con lo que muchos llaman la «democratización» de la tecnología de la información. Con una introducción cada vez mayor entre las masas, la significativa reducción del coste y el amplio abanico de aplicaciones disponibles, cualquiera tiene el poder de sacar el máximo partido a la tecnología existente, en muchos casos sin un céntimo de inversión inicial.

La democratización de la tecnología de la información no sólo ha afectado al espacio en la nube, sino también a los macrodatos en sí mismos. La adopción de tecnologías de macrodatos en código abierto crece a gran velocidad, y cada vez es más fácil efectuar análisis de datos con hardware de uso público o a buen precio.

Paralelamente, presenciamos una explosión de información generada a través de medios sociales, mensajes, emails, etc. Hoy en día individuos y organizaciones navegan por un laberinto de referencias cada vez más abundantes y por los que puede resultar difícil moverse, y más aún controlarlos y analizarlos.

De media, cada día se crean 2.500 millones de gigabytes, y cada mes se hacen 200 millones de tuits y se comparten 30.000 millones de contenidos en Facebook. Observando las proyecciones de las que disponemos, vemos que los datos creados en el año 2020 alcanzarán la imponente cifra de 43 billones de gigabytes, y 6.000 millones de personas dispondrán de teléfono móvil.

Pese a seguir en constante evolución, la computación en la nube y los macrodatos están demostrando ser la combinación ideal. Si alguien cree que los macrodatos no son más que una moda, que se lo piense dos veces. Tengamos en cuenta la enorme ventaja competitiva que supone poder acceder a un filón de fuentes de información, y el beneficio añadido que reporta la analítica.

En el mundo digital, la mayoría de cosas van abaratándose a gran velocidad. El precio de la banda ancha se ha reducido a menos de una centésima parte de lo que costaba hace quince años. Lo mismo ha pasado con el precio de la computación y el almacenaje, que se ha reducido drásticamente, alcanzando un coste prácticamente nulo. El coste medio de lanzar una empresa emergente también ha disminuido muy rápidamente, pasando de varios millones de dólares a decenas de miles en el caso de las *startups* de tecnología de internet. Con esta situación de abundancia informática (con costes casi nulos) y bajos costes de experimentación, nos adentramos rápidamente en un mundo impulsado por los datos, donde lo que adquiere valor son estos mismos. Los datos son lo que marca la diferencia, lo que da ventaja ante los competidores, la nueva moneda. Serán analizados, monitorizados, protegidos y cuestionados por los técnicos de éstos e IA, los nuevos gurús del ciberespacio, y afectarán a todas las industrias y a nosotros mismos como individuos.

2

El poder de las cosas y sus datos

Todos los científicos son científicos de datos; si no, no lo son.

JON CROWCROFT

La Internet de las Cosas

Cada vez son más los dispositivos conectados a internet, y se están haciendo más presentes cada día. El mundo conectado y rico en datos que están creando va a cambiar nuestras vidas de un modo que ni nos imaginamos. Internet primero conectó ordenadores, luego a la gente; ahora conectará cualquier cosa de nuestro entorno, dándole vida.

Si la idea de la Internet de las Cosas (IdC) te parece algo difícil de entender, piensa en esto:

Hubo una época, no hace mucho tiempo, en que la mayoría de personas no «entendía» qué era internet. Hoy en día, la mayoría lo entiende. ¿Qué es internet? Es una enorme cantidad de ordenadores hablándose unos a otros. Pero ahora que añadimos procesadores a todo tipo de dispositivos que antes eran tontos, la definición de «ordenador» también se ha ampliado.

Por un módico precio añadimos un minúsculo módem a uno de esos chips y —*voilà!*— conectividad para todo tipo de objetos. Y eso, por definición, es la Internet de las Cosas: artículos de uso común —desde bicicletas hasta relojes, termostatos, coches o gafas de realidad virtual— que antes no estaban conectados a internet, pero ahora sí.

Estos dispositivos IdC se usan para recopilar datos para su análisis, como un pulsómetro de esos que mandan información de tu trabajo físico al teléfono móvil. Eso permite automatizar los sistemas de un modo novedoso. Algunos coches conectados, por ejemplo, pueden comunicar con el termostato IdC de tu casa, y encender el aire acondicionado cuando el GPS informa de que te estás acercando a casa.

Pero en términos generales, si añades una conexión a internet a algo para poder controlarlo, obtener registros o ambas cosas, forma parte de la Internet de las Cosas.

Parece sencillo, ¿verdad? Lo es, y aportará enormes beneficios.

La IdC ya está aportando cambios significativos al modo en que cultivamos alimentos, usamos combustible, conducimos nuestros coches o controlamos nuestro estado de salud. La Monsanto Company y otras empresas agrícolas usan la IdC para hacer más fácil, más rápido y más eficiente el cultivo y la cosecha de alimentos.

En un artículo publicado el año pasado en *The New York Times*, Steve Lohr escribió: «La tercera fase [...] está empezando a arrancar, y supone explotar datos de muchas fuentes: sensores en las instalaciones agrícolas y en las plantas, imágenes por satélite y seguimiento de la meteorología. En un futuro próximo, el uso del agua y el fertilizante se medirá y se monitorizará con detalle, en algunos casos planta por planta».

Y Monsanto ya ha llegado a ese punto. La empresa utiliza sensores para recopilar datos en tiempo real con las cosechadoras, medir el volumen de las cosechas, la calidad de la tierra y la humedad, con el fin de aumentar la producción. También usan sensores para controlar la temperatura y la localización de las remesas de semillas, de modo que en caso de calor extremo puedan adaptar las rutas y así reducir las pérdidas de semillas.

Pero la IdC no sólo está cambiando el modo en que se cultivan los alimentos. General Electric (GE) lleva usando sensores IdC, programas y sus propias técnicas de análisis para ahorrar muchísimo dinero a sus clientes.

La tecnología operacional de la empresa le permite reducir el tiempo que pasan las plantas de gas natural licuado (GNL) fuera de servicio, señalando riesgos potenciales antes de que se conviertan en graves problemas. Con eso las empresas de GNL están ahorrando hasta 150 millones de dólares al año, tal como ha publicado la *MIT Sloan Management Review*.

GE empezó a avanzar hacia lo que se llama la Internet Industrial (básicamente, una subdivisión de la IdC), porque vio una ocasión de sacar beneficio de la recopilación y el análisis de datos sobre el uso del equipamiento industrial. La compañía actualmente considera que en 2020 los ingresos generados por sus soluciones de software alcanzarán los 15.000 millones de dólares.

A un nivel más personal, la IdC está cambiando nuestro modo de conducir, e incluso nuestra salud. Los coches semiautónomos (que pueden circular solos, en la mayoría de casos sólo por autopistas) ya están en las calles, y están abriendo el camino a una autonomía mucho mayor.

Recientemente Intel, BMW y Mobileye se asociaron para crear coches de empresa sin conductor en 2021. La empresa israelí Mobileye usa sensores instalados en los coches para enviar a la nube datos de localización con los que crear mapas, de

modo que otros coches puedan aprender de ellos. Lo llaman «Gestión de la Experiencia Viaria» (*Road Experience Management*; REM), y permitirá que los coches sin conductor sepan dónde se encuentran gracias a la información que comparten unos con otros.

Mobileye combinará la REM con sus chips EyeQ, ya disponibles, que procesan las imágenes de alrededor del coche para darle al vehículo una idea de su propia situación. Aproximadamente un 90 por ciento de los fabricantes de coches usan ya la tecnología de Mobileye, y la empresa espera que sus chips EyeQ de quinta generación sean los sensores del ordenador central de los coches completamente autónomos a partir de 2020.

Y no hay que pensar que los coches sin conductor son pura ciencia ficción. IHS Automotive calcula que el 10 por ciento de todos los vehículos ligeros que se venderán en 2035 serán completamente autónomos.

Ahora mismo, sobre todo nos imaginamos la tecnología IdC ayudándonos a mejorar la producción mundial de alimentos. Y toda ayuda es poca. La población mundial se ha duplicado en los últimos cincuenta años y se calcula que aumentará hasta llegar al menos a los 8.000 millones de personas en 2050. Por ejemplo, una empresa llamada Semios está aplicando redes de sensores que detectan la escarcha, la humedad de las hojas, la humedad del terreno y la presión de las plagas. Cuando la presión de las plagas alcanza un nivel crítico, el sistema distribuye unas dosis precisas de feromonas para alterar los ciclos de apareamiento de las especies invasoras.

También hay que insistir en la necesidad de hacer un uso óptimo del agua. En las regiones productoras de alimento donde hay escasez de agua, cuando se producen ciclos de sequías impredecibles, por ejemplo, es esencial dar una respuesta medida y bien argumentada. Los granjeros que se enfrentan a estas situaciones necesitan disponer de información precisa para reducir las pérdidas y los usos inadecuados al mínimo. Los sistemas inteligentes de riego, con sensores que registran la humedad del terreno y permiten que los granjeros valoren la necesidad en relación con la disponibilidad, están demostrando ser muy útiles.

Todos hemos oído hablar de la idea de usar etiquetas RFID en los envoltorios de los alimentos para que avisen a una nevera inteligente cuando la leche se haya pasado, o cuando quede poco helado. Entonces, la red inteligente de casa enviará una notificación *push* al *smartphone* del usuario. Pero la IdC también se está usando para controlar la salud del ganado. Un informe publicado en theguardian.com explica que un grupo de investigadores están usando datos recogidos de las vacas en libertad para diseñar algoritmos que permitan optimizar la producción de leche o de carne, o marcadores que avisen cuando una vaca enferma.

El síndrome del colapso de las colmenas (CCD por sus siglas en inglés), que ha afectado a poblaciones de abejas de toda América del Norte, es un motivo grave de

alarma. Sin nuestra principal especie polinizadora, desaparece el sector de la agricultura. Nadie conoce con seguridad el espectro completo de causas de este síndrome, pero sí conocemos algunos factores que lo propician.

Uno de los factores es un tipo de ácaro que ataca a las abejas, llamado *Varroa destructor*. Estos ácaros son fácilmente exterminables. Aumentando la temperatura a 42 °C estallan como granos de maíz, y esa temperatura no es ningún problema para las abejas. De modo que una empresa llamada Eltopia ideó un producto llamado provisionalmente «Mite Not». Es un circuito impreso en una película de bioplástico de maíz que recubren con cera de abeja, de modo que las abejas lo puedan usar para construir su panal.

El circuito funciona como sensor de temperatura, pero también como resistencia, y cuando se dan ciertas circunstancias en el panal, el sistema envía una señal al circuito para que caliente el panal, o una parte del panal. Es un ejemplo de exterminación de plagas de alta tecnología, sin pesticidas.

Por último, la IdC ha empezado a transformar nuestro estado de salud. Hoy en día los seguidores de actividad y los relojes inteligentes son los dispositivos tecnológicos portátiles que más atención están recibiendo, pero ya hay en el mercado dispositivos portátiles de uso médico u hospitalario.

Microsoft produce software que funciona con prendas y dispositivos portátiles y que permite que los pacientes hagan un seguimiento en tiempo real de la salud del paciente. Eso no sólo tiene la ventaja evidente de que ofrece acceso inmediato a la asistencia en caso de que al paciente le pase algo, sino que además reduce el tiempo empleado en diagnosticar los problemas de salud y da más tiempo para el tratamiento.

Por otra parte, Verily Life Sciences, de Google Alphabet, está desarrollando nanopartículas que pueden identificar el cáncer y otras enfermedades, y enviar sus datos a un dispositivo portátil situado en el brazo del usuario. La idea, aparentemente ambiciosa, es que estas partículas puedan llegar a detectar enfermedades en fases tempranas.

Estos dispositivos y este software podrían ahorrar enormes cantidades de dinero, además de mejorar nuestra salud: GE cree que sólo aumentando un 1 por ciento la eficiencia de la asistencia sanitaria se ahorrarían 63 millones de dólares cada año.

Año tras año, a medida que la Internet de las Cosas va aumentando en sofisticación, iremos viendo gran cantidad de referencias sobre nuestro entorno (visuales, de localización, de audio, gráficos, 3D, renderizaciones, temperatura, etc.) en tiempo real, y eso supone una enorme cantidad de datos. Por ejemplo, un coche que se conduzca solo genera unos 10 gigabytes de datos por cada milla recorrida (1,6 km

aprox.), y éstos tienen que procesarse en menos de 10 milisegundos, tiempo necesario para que los sistemas anticolidión puedan intervenir y evitar un choque.

Enviar esa enorme cantidad de registros de los sensores y de vídeo a internet y procesarla en tiempo real será cada vez más difícil. Las infraestructuras existentes de internet no están diseñadas desde un principio para que puedan gestionar la gran revolución de los datos de la IdC (sensores, realidad virtual (RV)/realidad aumentada (RA), drones, vehículos autónomos, etc.). De hecho, ya hemos estado en ese punto, en ese momento en que todo puede cambiar en un instante.

Cómo acabar con internet

Era el verano de 1997, y yo empezaba mi doctorado en el Swiss Federal Institute of Technology de Lausana. Era el mismo día en que Lady Di moriría en el túnel del Pont de l'Alma de París, poco después de que Microsoft anunciara una histórica inversión de 150 millones de dólares en su histórico rival, Apple, de que el cometa Hale-Bopp alcanzara el punto más cercano a la Tierra de los últimos 2.000 años. Yo no sabía una palabra de francés y había pasado el verano tomando clases para poder sobrevivir en una de las universidades más duras del mundo, situada en un cantón suizo francófono, y completar mis estudios de doctorado. En aquellos días estaban cambiando muchas cosas, tanto para mí como en el mundo, y estaban a punto de cambiar muchas más.

Nada más llegar a mi despacho de la universidad, el lunes por la mañana, me encontré un artículo sobre el escritorio que decía: «¿Cómo acabar con internet?» acompañado de una nota: «Soy Ernst, hablamos más tarde». En mi primer día como estudiante de doctorado, mi tutor de tesis, Ernst Biersach, me había dejado un estudio reciente de un profesor del Lawrence Berkeley Laboratory, uno de los pioneros de internet (inventó el protocolo que evita la congestión de la red) para que lo leyera y me sirviera de inspiración en mi recorrido hacia el doctorado.

Empecé a leer el artículo y, justo después de aquella misteriosa pregunta encontré una respuesta aún más inquietante. A la pregunta «¿Cómo acabar con internet» le seguía: «Internet fue creada para que pudiera sobrevivir a una guerra nuclear pero, ¿cómo puedes acabar con ella? [...] Muy fácil: ¡Inventa la web!»

¿Cómo?! ¿La web podía acabar con internet? A diferencia de lo que piensan muchos, la web e internet son dos cosas diferentes. Internet es la red que proporciona la conectividad y que conecta ordenadores y redes de todo el mundo para que podamos viajar de unos a otros. La web vive «encima» de la red, y se compone de contenidos, páginas, redes sociales y servicios. Entre la web e internet hay una separación de casi treinta años.

Cuando empecé mi recorrido hacia el doctorado, la web acababa de ser creada en el acelerador de partículas del CERN, en Ginebra, cerca de donde estudiaba yo, e

internet, que había sido diseñada como una red de redes capaz de soportar grandes interrupciones por ataques militares, empezaba a tener dificultades para proporcionar sus contenidos web (imágenes, páginas, actualizaciones de seguridad, y más tarde vídeos) a un creciente número de usuarios.

El problema era que internet había sido diseñada para enviar mensajes entre ordenadores de todo el mundo pero, a diferencia de las redes de televisión o de radio, no para distribuir una enorme cantidad de contenido multimedia. Y ése era el desafío que tenía que resolver en los años siguientes para conseguir mi doctorado: tenía que asegurarme de que internet sobreviviera a la web.

Durante muchos años, internet había sido un banco de pruebas y un terreno de juegos para científicos como yo, antes de que se nos uniera el resto del mundo. Hoy en día es una red muy compleja y tiene una propiedad antifragilidad por la que no hace falta que todo funcione bien en todo momento para que siga funcionando. De hecho, falla y se repara constantemente para asegurar su disponibilidad en todo momento. No obstante, no siempre fue así.

Con cada día que pasaba, internet crecía y se congestionaba cada vez más. A finales de diciembre de 1996, internet se componía de más de 100.000 redes que conectaban más de 13 millones de ordenadores a decenas de millones de usuarios en todo el mundo. En enero de 1993 había cincuenta servidores web en todo el mundo; en octubre de 1993 había más de quinientos; a finales de los años 1990 se produjo la burbuja puntocom, y para 1997 ya había 100.000 sitios web. Hoy en día la web cuenta con más de 45.000 millones de páginas y entradas de datos.

La World Wide Web se estaba convirtiendo rápidamente en un sistema de diseminación de datos a escala global. El éxito sin precedentes de la web provocó una sobrecarga de tráfico en las fuentes de datos y las rutas de la red. Desde entonces se han observado repetidamente «puntos calientes» de sobrecarga, embotellamientos conocidos como hot spots que a finales de los años 1990 aparecían a diario. Para poder hacer frente a aquel desafío, había mucha gente en el emergente sector comercial de internet que esperaban que científicos como yo encontráramos la solución antes de que internet llegara a colapsarse.

De hecho, se colapsó. En 1999, el primer desfile de ropa interior de Victoria's Secret en la web generó tal demanda que internet se bloqueó en Dallas y otras zonas de Texas. ¡Victoria's Secret reventó internet! Sí, eso fue una gran cosa para el departamento de marketing de Victoria's Secret, pero una mala noticia para la web, ya que demostraba que internet no estaba preparada para distribuir una cantidad masiva de contenidos al mismo tiempo. Y durante los terribles atentados terroristas del 11 de septiembre, en pocos minutos todo el planeta acudió a los medios estadounidenses en línea en busca de más información. Con el aumento de la demanda de información de usuarios de todo el mundo, la mayoría de sitios web estadounidenses, como la de la CNN, no daban abasto. Algunos sitios web tuvieron que suspender el servicio, y muchos tuvieron que recurrir a sitios estáticos sin

imágenes, sólo con texto: era como ver la televisión en blanco y negro. Internet estaba fracasando como medio de comunicación mundial cuando más necesaria era.

En los años 1970 la comunidad de internet pensaba que necesitarían conectar todos los ordenadores del mundo, pero consideraban que sólo habría unos centenares de miles. Lo que sorprendió a todo el mundo fue la aparición del ordenador personal, porque de repente internet tenía que conectar no centenares de miles de ordenadores, sino centenares de millones. A principios de los años 1980 el gran problema era la gran escala de la infraestructura necesaria para conectar todos aquellos ordenadores. Los expertos en internet tenían que rehacer los protocolos para redimensionar la red, los organismos de normalización para redimensionarlos, etc. Pero en los años 1990 y principios de los años 2000 la misión era redimensionar internet para que pudiera distribuir el contenido web y los datos que quería la gente (páginas web y contenidos multimedia), no conectar ordenadores.

En 1995, Bob Metcalfe, inventor del aclamado protocolo Ethernet que conecta los ordenadores a internet en cada domicilio y empresa, hizo una predicción que se haría famosa. Predijo que internet sufriría un «colapso catastrófico»; y prometió comerse sus palabras si se equivocaba. De hecho, durante el discurso inaugural de la sexta World Wide Web Conference, en 1997, cogió una copia impresa del artículo en que predecía el colapso, la metió en una batidora con algo de líquido y luego se bebió la papilla resultante. Eso fue después de que intentara comerse sus palabras en forma de enorme pastel, pero el público protestó vehementemente, de modo que acabó comiéndose la versión líquida de su artículo.

Quizá Bob no tuviera razón en cuanto al momento y la magnitud del colapso de internet, pero desde luego la web estaba poniendo en riesgo la red, e internet precisaba un replanteamiento y una reorganización a fondo para poder estar seguros de que la predicción de Bob no se hacía realidad y para redimensionar la distribución de contenido web de alta calidad a cientos de millones de personas. ¿Cómo se hizo?

Casi sin darme cuenta, yo mismo me encontré inmerso durante más de una década en un viaje fascinante para reparar internet, diseñar nuevas arquitecturas de red que pudieran soportar la explosión de la World Wide Web y una gran cantidad de registros y algoritmos de internet que ayudaran a comprenderla y mejorarla. Para hacerlo tuvimos que construir una «nueva» internet virtual sobre la ya existente, con millones de ordenadores y repetidores de datos web nuevos, situados en lugares estratégicos del planeta, cerca de ti, en la puerta de tu casa. Así nos asegurábamos de que la red pudiera pasar de una velocidad de marcación por teléfono a una rápida distribución de datos, haciendo posible el acceso a vídeos e información de mayor calidad, la protección contra amenazas de seguridad y ciberataques más sofisticados, y un mayor tiempo de conexión a la red, tanto desde el ordenador como, más tarde, desde los teléfonos móviles.

Dedicamos muchos años a analizar conjuntos de macrodatos, entre ellos los de rendimiento de internet desde cada nódulo, desde el punto de vista de millones de usuarios y sitios web. Y usamos esos datos y algoritmos para rediseñar internet de modo que pudiera distribuir una enorme cantidad de información. Y yo mismo — considerado el joven revolucionario entre los «gurús de internet», ya que usaba novedosos diseños de comunicación *peer-to-peer* (P2P) que rompían con los diseños tradicionales existentes, más jerárquicos—, desde luego estaba fascinado.

Una distribución de datos más rápida

La web la inventó Sir Tim Berners Lee en la Organización Europea de Investigación Nuclear del CERN de Ginebra a finales de los años 1980. Internet fue un invento del programa ARPANET de la Agencia de Proyectos de Investigación Avanzados de Defensa de Estados Unidos en cooperación con varias universidades (UCLA, Stanford, UCL) a finales de los años 1960. La web se compone de una gran cantidad de centros de procesamiento de registros (CPD) que almacenan el contenido web por el que navegamos, las páginas web, imágenes, vídeos, televisión, y que también registra las interacciones de nuestros datos y nuestro rastro digital. Estos centros de procesamiento de datos están conectados a internet a través de una serie de cables que transportan la información hasta tu casa o tu dispositivo móvil.

El término «internet» hacer referencia a la red de redes, a una red que conecta a millones de redes privadas, públicas, académicas, empresariales y gubernamentales, de ámbito local o global, unidas por una amplia gama de conexiones electrónicas, *wireless* u ópticas y que te da acceso a una gran variedad de servicios y recursos web. Es la red de base por la que viajan los datos.

Básicamente, internet se compone de una serie de redes de acceso local (fijas o inalámbricas) a nivel urbano, interconectadas con un número mayor de redes a nivel nacional y continental (redes de interconexión) que te conectan con centros de procesamiento de datos (o con la nube), que alojan una gran cantidad de ordenadores, en muchos casos en puntos fríos y remotos del mundo, para disponer de espacio y ahorrar en refrigeración.

Lo que hace que internet sea tan especial es que es un sistema que permite que diferentes redes de ordenadores se comuniquen usando una serie de normas estandarizadas y se comprendan entre sí. Es verdaderamente un sistema global. Cables que atraviesan países y océanos, que traspasan fronteras y que comunican algunos de los lugares más remotos de la Tierra con todo los demás. E internet sigue creciendo. Cada día hay más ordenadores conectados, aparecen y desaparecen redes, y numerosas organizaciones y empresas trabajan para ampliar el acceso a internet en países que aún no están conectados. Internet no es algo físico y sólido; es más bien un concepto variable localizado en todas partes.

Pero en el momento en que iniciaba mi trabajo de investigación científica nos dimos cuenta de que había un problema: internet no podría soportar la cantidad de datos que la web podía proporcionar, especialmente en momentos destacados y eventos críticos en que todo el mundo visitaba los mismos lugares de la red. De hecho, si sumamos toda la velocidad a la que pueden enviar información los CPD a la red de internet y la comparamos con la velocidad máxima a la que la gente puede descargarse cosas de internet, hay una gran diferencia. La nube, donde se encuentra la información, sólo puede proporcionar datos a más o menos un 2 por ciento de la velocidad máxima a la que pueden recibirla los terminales.

De entre todos los que trabajábamos en este sentido, como los científicos que, como yo, buscaban mejorar la velocidad de internet, había una empresa que había aparecido en pleno nacimiento de la web, una *startup* que acabaría proponiendo una solución a la explosión de datos de la web: Akamai.

Artistas, magos, científicos como yo, chefs y diseñadores tienen en común el hecho de que comparten la mayor parte de su trabajo en foros abiertos en los que pueden debatir, potenciar el trabajo de los demás, obtener *feedback* y crecer como comunidad, en beneficio de todos. En uno de estos eventos, durante mis estudios de doctorado, conocí a Daniel Lewin, colega estadounidense que también trabajaba en la mejora de la web y en el diseño de algoritmos de distribución de datos que más adelante se volverían cruciales para la red. Estos algoritmos permitían que un gran número de servidores compartieran la carga con muy poca información de estado, repartiéndola de un modo muy efectivo. Daniel Lewin estaba haciendo sus estudios de doctorado en el Massachusetts Institute of Technology (MIT) en 1996. En aquel tiempo, Lewin y su tutor de tesis, el profesor F. Thomson Leighton, crearon Akamai.

Akamai desplegó 95.000 servidores en 72 países, y con 2.200 empleados, la empresa proporciona del 15 al 30 por ciento de todo el tráfico de la web en el mundo. Hoy en día, las innovaciones de Akamai permiten que millones de usuarios puedan ver vídeos en *streaming* a la vez, y mantiene los sitios web de noticias en línea durante las crisis globales, cuando los usuarios se lanzan en busca de la última información. También proporciona una barrera de seguridad a internet, evitando ataques de denegación de servicio, ataques de virus y *malware*, problemas de ciberseguridad y dando protección en importantes procesos de votación, en los que internet cada vez tiene un papel más crucial.

Internet necesitaba un modo mejor para ubicar de un modo instantáneo las enormes cantidades de datos cambiantes almacenados en ordenadores de todo el mundo y enviarlos a cualquier usuario, en cualquier lugar. E —igual de importante— el sistema también debía poder «redimensionarse», lo que significaba que pudiera funcionar aunque fuera utilizado por un número mucho mayor de personas. Es lo que posibilitó los avanzados servicios de internet que usamos hoy en día. La tecnología de Akamai gestionó el problema con habilidad; sitios web que de otro modo se

habrían venido abajo con la fuerte demanda funcionaban ahora muy bien. Pero las cosas no eran tan sencillas.

El 11 de septiembre de 2001 yo vivía en Palo Alto (California). A primera hora de la mañana recibí una llamada de teléfono de mi familia en Europa. Unos terroristas habían atacado Nueva York usando aviones. Yo fui corriendo a internet para obtener más información. Daniel Lewin, el estudiante de doctorado que había coincidido en una conferencia para discutir sobre nuestro trabajo para mejorar la web, había dejado el MIT y había fundado Akamai con su profesor. Daniel Lewin también había servido cuatro años en las Fuerzas de Defensa Israelíes como oficial del Sayeret Matkal, una de las unidades de fuerzas especiales del ejército.

Desgraciadamente, Daniel Lewin acabó siendo la primera víctima de los ataques terroristas contra Estados Unidos que desencadenaron más de una década de guerra. Lewin se encontraba a bordo del vuelo nº 11 que impactó contra la Torre Norte del World Trade Center. Volaba de Boston a Los Ángeles por negocios relacionados con Akamai. Según parece, Lewin fue apuñalado durante el secuestro del vuelo nº 11 de American Airlines ese 11 de septiembre. Un informe de la Administración Federal de Aviación de Estados Unidos sugiere que quizá lo apuñalara Satam al-Suqami cuando intentó frustrar el secuestro. Lewin iba sentado en clase preferente. Ocupaba la plaza 9B, cerca de los secuestradores Mohamed Atta, Abduaziz al-Omari y al Suqami. Fue apuñalado por uno de ellos, probablemente por Satam al Suqami, que estaba sentado justo detrás de él. La comisión de investigación de los atentados especula que quizá ocurriera cuando Lewin intentó enfrentarse a uno de los secuestradores que tenía delante, ignorando que al Suqami estaba sentado detrás. Mientras tanto, sus colegas de Akamai, en Boston, vivían uno de los días más decisivos de la empresa, poniendo a prueba los algoritmos diseñados para mejorar internet después de años de investigación.

Poco después de las 9.00 horas del día de los atentados, un grupo de empleados del Centro de Control de Operaciones de la Red de Akamai Technologies, que tenía su sede en Cambridge, habían observado un aumento inusual de tráfico en internet. Cuando se enteraron de que era debido a la noticia de que una pequeña avioneta se había estrellado contra el World Trade Center, se alarmaron, y no sin motivo. Muy pronto, no obstante, se hizo evidente que no había nada en aquel incidente que fuera pequeño. Después del impacto del segundo avión, prácticamente cualquiera que tuviera una conexión a internet estaba conectándose para enterarse de qué había pasado.

Eso supuso una gran presión para sitios web de muchos de los clientes de Akamai, entre ellos la ABC.com, *The Washington Post* y otros muchos medios de comunicación, así como para los sitios web de la Cruz Roja, el FBI, American Airlines (propietaria del primer avión que impactó contra las torres, el vuelo nº 11, y

del nº 77, que impactó en el Pentágono). El pico del 11/9 fue el equivalente en la web a una inundación de cien años de duración.

El alud de solicitudes también afectó a los servidores de nombres de dominios de internet (DNS), que actúan como guía de teléfonos de internet, facilitando el acceso a páginas web específicas. Son los encargados de asegurarse de que acabas en el lugar correcto cada vez que escribes el nombre de un sitio web en tu navegador. Los hackers a veces atacan a los proveedores de DNS para sabotear los sitios web a los que redirigen. Eso sigue siendo un gran problema hoy en día, y un recurso habitual de los hackers para atacar la red. Era responsabilidad de Akamai procurar que las empresas que pagaban por sus servicios se mantuvieran conectadas, y solucionar los problemas a los que se enfrentaba internet con el éxito de la web.

Diecisiete años es una eternidad, en tiempo de internet. En 2001 no teníamos Twitter, Facebook ni Instagram. Aun así, cuando cayeron las Torres Gemelas, los estadounidenses se dirigieron en masa a la red para buscar información sobre una de las noticias más importantes de la década. A pesar de lo limitados que eran muchos servicios —aún faltaban cuatro años para la llegada de los vídeos generados por los usuarios al estilo Youtube—, el 11 de septiembre produjo una cantidad de datos increíble.

Antes de que las Torres Gemelas se vinieran abajo, la empresa de direccionamiento del tráfico sólo había movido, como mucho, 6 gigabits de datos por segundo. Entonces la MSNBC decidió emitir su programa en directo por internet, lo que elevó el uso de éstos de la emisora por cable muy por encima de lo que había registrado nunca Akamai. La MSNBC había contratado a Akamai justo para que resolviera ese problema, y con el aumento del uso de aquéllos, Akamai observó que el sitio web de la MSNBC movía 12,5 gigabits de datos por segundo. (Para que nos hagamos una idea, hoy en día toda Nueva Zelanda hace un uso de banda ancha equivalente a 120 gigabits por segundo en horas pico).

Para algunos, internet se volvió lentísima. Cuando yo me conecté para informarme, intenté buscar noticias de Yahoo y de la CNN, pero sus sitios web se cargaban de forma intermitente. De hecho, el sitio web de la CNN tuvo que pasar a usar páginas de texto, informativas pero muy básicas.

La infraestructura de internet no fue el único obstáculo. Aquello pilló desprevenidos a todos los sitios web. El FBI, que había colgado fotografías de los terroristas del 11 de septiembre a pocas horas de los atentados, también tenía problemas para mantenerse a flote con todo aquel tráfico suplementario.

Las empresas de internet han aprendido mucho en los diecisiete años que han pasado desde entonces. Ahora se adaptan de forma rutinaria a los grandes acontecimientos noticiables ante la posibilidad de que la gente acuda en masa a la web para enterarse de qué ha sucedido, y para participar en el proceso de recopilación de información.

Se trate de una boda real, de las Olimpiadas o del Black Friday, siempre va a haber algún evento, programado o espontáneo, que generará una carga exorbitante para la que no estaba preparada internet.

Para solucionarlo ha habido que hacer muchos análisis de datos e investigación algorítmica. De hecho, uno de los problemas más importantes que hemos tenido que solucionar para hacer que la web vaya rápida fue el de localizar el servidor web más rápido que puede conectar con un usuario, ya que encontrar el servidor más rápido es esencial para tener una experiencia rápida y positiva.

Las páginas web y los archivos de vídeo se almacenan en grandes centros de procesamiento de datos llenos de ordenadores repartidos por el mundo. La intuición nos dice que cuanto más haya que viajar por internet para encontrar un servidor web que tenga el vídeo que se busca, más posible es que la calidad percibida por el usuario sea mala, ya que las probabilidades de encontrarse con un nodo congestionado aumentan. La congestión de la red aumenta con la distancia, y los servidores más rápidos son los que tienes cerca. Por ejemplo, para un usuario en Madrid, un servidor de Londres será más rápido que uno en Mountain View (California).

Cabe pensar que encontrar el servidor más rápido es tan sencillo como trazar una línea sobre el mapa y encontrar el que esté más cerca geográficamente. No obstante, el problema en realidad es mucho más complicado, y aquí la intuición falla. Internet tiene algunos caminos «ocultos» por los que se pueden hallar servidores más rápidos que los que encontrarían trazando una línea recta sobre el mapa. En internet, el camino más rápido no siempre es el más corto; puede ser el más largo, tal como ocurre en muchas otras situaciones de la vida. ¡Resulta que internet es una red curva! Del mismo modo que el universo es curvo, o que la Tierra es curva, internet también es curva y tiene caminos que desafían a la intuición y que hacen que algunos servidores lejanos sean más rápidos que otros situados más cerca.

Cuando haces un vuelo transatlántico, ¿por qué no va recto el avión? ¿Por qué traza una curva y se acerca a Groenlandia? Cabría pensar que la ruta más corta sería una línea recta, pongamos entre Nueva York y Londres. Pero el avión traza una curva. Lo que aparentemente es el camino más lento es, de hecho, una vía más rápida, siguiendo rutas «más largas» no evidentes a primera vista. Algo parecido pasa en internet. Hay caminos más largos que son más rápidos que los cortos, e internet también parece tener una extraña forma curva, elíptica o hiperbólica.

Ésta es una sorprendente conclusión que surgió a partir de numerosos análisis de datos y que llevó décadas descubrir, pero es de vital importancia para el buen rendimiento de internet. Esta propiedad es crítica para el funcionamiento de la red sin un sistema de coordinación central que conecte todas las redes del mundo, y para permitir el redimensionamiento de la web, redirigiendo los picos de sobrecarga, como el del 11 de septiembre, al mejor servidor de cada momento.

Los macrodatos ya juegan un papel decisivo en nuestras vidas. Con el crecimiento de la IdC, esa influencia no dejará de crecer, ya que la Internet de las Cosas genera más datos, *feeds* de vídeo e imágenes que nunca. Vivimos en el zénit de una revolución digital nunca vista desde la aparición de internet en nuestras vidas. Sólo es cuestión de tiempo hasta que los datos digitales se conviertan en componente de prácticamente todo lo que hacemos. Y desde el momento en que se creó internet éstos se han usado para hacer que la gente pueda disfrutar de una internet más rápida y mejor, incluso antes de la llegada de la IdC.

Los científicos calculan que nos encontramos en la cúspide de la cuarta revolución industrial, en la que la inteligencia artificial, la robótica y la Internet de las Cosas cambian el modo en que hacemos las cosas. Primero la mecanización y el vapor transformaron nuestras vidas; luego fue la electricidad; recientemente fue la digitalización e internet; la próxima revolución será la de la inteligencia artificial, la robótica y la Internet de las Cosas.

La IdC a la velocidad de la luz

El próximo desafío para internet es el de asegurarse de que todos los dispositivos IdC puedan comunicarse entre sí como si estuviéramos en el mismo lugar. Eso será necesario para posibilitar la existencia de espacios de realidad virtual o de realidad aumentada en los que nos dé la impresión de tener a nuestros seres queridos o más próximos en la misma estancia que nosotros, o para que podamos fusionar la realidad virtual y la vida real, para que los desarrolladores de programas puedan crear imágenes en su interior que se fundan con la vida real. Con la realidad virtual y la realidad aumentada los usuarios pueden interactuar con los contenidos virtuales del mundo real y pueden distinguir los unos de los otros.

La realidad virtual y aumentada serán muy importantes para las empresas. Facebook compró Oculus Rift por 2.000 millones de dólares aunque la empresa no tenía clientes y sólo tenía dieciocho meses de vida. Con ello Facebook demostró que éste será un nuevo modo de interactuar con los ordenadores y con las personas. Con la realidad aumentada podrás pintar el mundo real. Y no son los únicos: Microsoft Hololense permitirá una experiencia similar para aprender, adiestrarse a distancia y experimentar nuevas realidades. La realidad virtual y la realidad aumentada requerirán una gran anchura de banda y mucha velocidad para poder aportarte la experiencia que buscas. Pero, sobre todo, necesitarán períodos de latencia y retrasos mínimos. Todo tendrá que responder muy rápidamente. Si no, la experiencia será muy pobre. ¡E incluso puede que te marees!

Hoy en día internet es tan rápida que la bahía de San Francisco puede conectar con Nueva York en menos tiempo del que se tarda en acabar esta frase. Eso es una velocidad de vértigo, pero no lo suficiente para los músicos, por ejemplo, que sueñan con el día en que podrán conectar los instrumentos a internet para que sus notas viajen a la velocidad de la luz. Así todo el mundo podría tocar en una única formación musical.

Los retrasos tienen un efecto devastador. Para solucionar este problema, la nube tiene que acercarse más a los consumidores. ¿Treinta milisegundos? Eso es un eco. Una eternidad. La esencia de la música es la comunicación en común. Y eso depende de un *feedback* instantáneo. Así que todas las esperanzas están puestas en un intento por superar el límite de velocidad, lanzado en octubre de 2014 por un equipo encabezado por Brighten Godfrey, investigador informático de la Universidad de Illinois y sus colegas de la Universidad de Duke.

Su misión —«trabajar en red a la velocidad de la luz»— desafía a los investigadores de redes informáticas a que creen una internet que alcance el límite físico del universo. Ellos se imaginan a un par de músicos de Oahu, en Hawái, tocando un cuarteto de Beethoven al unísono con otros dos en la Antártida.

Y éste es el problema: el retraso —o «latencia»— de una señal al recorrer una red. Los paquetes de datos viajan por internet unas 10 veces más despacio que la velocidad de la luz, a veces hasta 100 veces.

¿Esa conexión San Francisco-Nueva York? En la red informática típica, tarda un segundo más o menos; o incluso más. Si viajara a la velocidad de la luz, sólo tardaría 27 milisegundos (un milisegundo es una milésima de segundo).

Esa lentitud relativa resulta frustrante para los músicos, porque el oído humano percibe dos sonidos como simultáneos sólo si los oye a menos de 20 milisegundos de distancia. De modo que, aunque un grupo de músicos conectados en red en la misma ciudad podrían tocar juntos, resulta mucho más difícil si están separados por grandes distancias. (Incluso a la velocidad de la luz, unos 300.000 kilómetros por segundo, una melodía bidireccional tocada por dos músicos situados en extremos opuestos del planeta presentaría un retraso de 133 milisegundos).

Los retrasos también son un gran problema para las empresas de internet, que pierden millones de dólares cuando se reduce la velocidad. En el caso de Google, un retraso adicional de 400 milisegundos en la respuesta a las búsquedas reduce el volumen de búsquedas en un 0,74 por ciento.

En los últimos veinticinco años, la velocidad de internet ha aumentado a un ritmo asombroso; aun así, los investigadores están haciendo que sea aún más rápida. El doctor Brighten Godfrey, profesor auxiliar de la Universidad de Illinois en

Urbana-Champaign, trabaja para que las cosas se puedan conectar por internet a una velocidad lo más próxima posible a la de la luz, barrera física insuperable.

Hoy en día, cuando las personas y las máquinas interactúan en la red, hasta para enviar un mensaje breve y obtener una respuesta hace falta un tiempo significativo. Eso es lo que llamamos latencia de la red. «La latencia es el gran reto al que se enfrentan nuestras experiencias interactivas en internet, tanto entre personas como entre ordenadores —señala el profesor Godfrey—. Y hoy en día cada milisegundo importa».

La latencia de la red tiene un impacto increíble en la posibilidad de conectar cosas. Sin embargo, hoy en día internet suele funcionar a una velocidad 30 veces mayor —y en muchos casos hasta 100 veces mayor— de lo que debería imponer el límite de la velocidad de la luz.

Y la mejora de la latencia de internet puede tener un profundo efecto en cómo usan internet las cosas. Internet debería operar a la velocidad de la luz. El tiempo que se tarda en enviar datos de un punto A a un punto B del mundo debería acercarse lo máximo posible a la velocidad a la que viaja la luz. Pero si medimos la rapidez de internet en la actualidad nos daremos cuenta de que para alcanzar un dato cualquiera tardamos unas veintitrés veces más de lo que se tardaría a la velocidad de la luz. Para solucionar este problema hay que recurrir a numerosos análisis de datos que permitan mejorar los protocolos de internet, y situar servidores cerca de cada usuario del mundo, reduciendo así el tiempo que tarda en viajar la información.

Las máquinas de operaciones financieras rápidas como la que describe el libro *Flash Boys* requiere unas conexiones de internet superrápidas. Las páginas de comercio electrónico como Amazon, Google o Facebook han demostrado repetidamente que reduciendo una décima parte el tiempo que se tarda en visualizar una página se traduce en aumentos de beneficios mesurables en sus páginas comerciales (y un 1 por ciento más de ingresos aproximadamente).

El acceso rápido a los datos es tan importante que las empresas han empezado a construir sus propias redes y a reducir las distancias. Es el caso de Spread Networks, empresa de operaciones en bolsa que hace unos años completó un nuevo túnel que atraviesa los montes Allegheny, en Pensilvania. No obstante, el túnel de Spread no estaba pensado para el paso de pasajeros o de mercancías; era para un cable de fibra óptica que ahorraría tres milisegundos —tres milésimas de segundo— en las comunicaciones entre los mercados de Chicago y la bolsa de Nueva York.

¿A quién le importan tres milisegundos? La respuesta es a las máquinas de comercio de alta frecuencia, que hacen dinero comprando o vendiendo acciones una fracción de segundo antes que los demás. No es de extrañar que Michael Lewis inicie el relato de su libro *Flash Boys*, gran éxito de ventas sobre la polémica generada por el comercio de alta frecuencia, con la historia del túnel de Spread Networks.

La historia de *Flash Boys* se centra en varias personas, entre ellas Sergey Aleynikov, que fue programador de Goldman Sachs, y Bradley Katsuyama, fundador

de Investors' Exchange (IEX). Flash Boys empieza describiendo la construcción de la línea secreta de 1.330 kilómetros de cable, para la que hacía falta un trazado lo más recto posible, atravesando montañas y pasando bajo los ríos, desde Chicago hasta Nueva Jersey, para reducir el tiempo de transmisión de los datos... ¡de 17 a 13 milisegundos! Este proyecto de 300 millones de dólares estaba diseñado para conectar los mercados financieros de Chicago y Nueva York, donde cabía la posibilidad de sacar beneficios con una inversión ventajosa realizada con un margen de milisegundos.

Recientemente la velocidad de los datos en los mercados financieros se ha convertido en un asunto de gran importancia: cuanto más rápido viajen éstos, mayor es el beneficio. Los expertos afirman que el acceso a estos cables de fibra óptica, al igual que otras tecnologías, dan al mercado de valores la posibilidad de obtener información más rápida, volviéndose así más eficientes. En los últimos años, el comercio electrónico ha sustituido al parque de la bolsa, con sus *brokers* que gritaban, daban golpes al teléfono y repasaban, histéricos, la cinta de la teleimpresora, y ese cambio ha tenido un impacto en el mercado.

Con los avances tecnológicos, los valores actualmente se compran y se venden digitalmente, en servidores informáticos, en edificios a menudo anónimos —pero muy vigilados—, generalmente a kilómetros de los epicentros históricos de las finanzas, lo que significa que los tipos audaces con trajes elegantes representados en películas como *El lobo de Wall Street* han dejado de ser los reyes de las finanzas.

Los programadores informáticos se han hecho con la corona gracias a los códigos que producen, capaces de ejecutar operaciones miles de veces más rápido que cualquier humano. En el centro están los operadores de operaciones de alta frecuencia (HFT por sus siglas en inglés), que instalan conexiones de datos ultrarrápidas de fibra óptica entre sus sistemas y los mercados de valores modernos, lo que les da una minúscula ventaja de velocidad sobre sus rivales. Esta ventaja, aunque sea de milisegundos, les permite ver las órdenes de otros compradores antes de que se ejecuten.

Básicamente, usan esta ventaja temporal para comprar las acciones antes de que se procese la primera orden, y para vendérsela al comprador interesado a un precio ligeramente más alto, en un proceso conocido como *front-running*. Esto significa que esos milisegundos valen millones.

La velocidad de los datos se ha convertido en algo tan importante que actualmente las compañías de inversión en bolsa invierten en redes de propiedad, como enlaces de radio (por microondas) entre Chicago y Nueva Jersey, que siguen una ruta aún más recta que el cable de 1.330 kilómetros de Spread Networks (ya que las microondas siempre siguen una trayectoria recta, mientras que los cables, por su propia naturaleza, deben rodear barreras físicas). La nueva ruta también aprovecha la velocidad superior a la que viajan las señales por el aire (en comparación con la velocidad de transmisión de la fibra de vidrio, que frena la luz). Con estas dos

ventajas, este nuevo enlace ha rebajado en 4,5 milisegundos la velocidad del cable de Spread Networks, y eso puede marcar una gran diferencia en el mercado de valores.

Pero ésa no es más que una aplicación posible. Hay otras aplicaciones interactivas —como las apuestas, la navegación en la web, la cirugía remota, el aprendizaje a distancia, la conducción autónoma o la colaboración musical— que se beneficiarían tremendamente con la mejora en la velocidad de internet.

Los humanos, con nuestro sistema de percepción visual, no podemos distinguir un evento que ocurre inmediatamente después de apretar un botón de uno ocurrido 30 milisegundos después de haber apretado un botón. Si se consigue que internet sea 30 milisegundos más rápida, a los seres humanos nos parecerá instantánea, desde el punto de vista visual. Eso significa poner un servidor de internet a menos de 900 kilómetros de cualquier ser humano del planeta, o a menos de 300 kilómetros si es necesario que el retraso no sea superior a los 10 milisegundos.

La tecnología ya está conectando cosas como instrumentos musicales por el mundo, y transformando el concepto de creación musical. Se pueden tomar clases y realizar audiciones por YouTube. Pero la colaboración en tiempo real sigue siendo la última frontera. Recientemente, en un concierto organizado por la Universidad de Stanford, se hizo público el compromiso de que la velocidad de internet serviría para construir puentes culturales a través de la música. En esta ocasión, el músico y científico Chris Chafe utilizó la conexión de alta velocidad de la universidad, Internet2, para poner en contacto a trece músicos que tocaban en Stanford, en la UC Santa Barbara, en la Virginia Tech y en la Universidad de Guanajuato, en México.

Los músicos, que se oían como si estuvieran tocando en la misma sala de conciertos, entretejieron una interpretación clásica, folclórica y electrónica llamada *Imagining the Universe*. Las notas del contrabajo y el piano de Virginia Tech tardaron 46 milisegundos en llegar a Stanford, igual que el violonchelo de Chafe. El tiempo que empleó el sonido de la flauta de México fue similar. Santa Bárbara, estaba más cerca, a sólo 7 milisegundos de distancia.

Hace treinta años no habríamos podido soñar siquiera que llegaríamos a este punto, pero hoy intentamos llegar más lejos, para unir de verdad todo el mundo y todas sus cosas.

La Internet de las Cosas y los datos

Internet primero conectó los ordenadores; luego conectó a las personas, y ahora está conectando las cosas. Y con «cosas» me refiero a todas las cosas físicas que nos rodean. Coches, lavadoras, ropa, sensores de las fábricas, alimentos en los supermercados, muebles, semáforos, etc. Eso es lo que se llama la Internet de las Cosas (*Internet of Things*, IdC), y lo que se espera que sea la nueva fase de crecimiento en internet.

Si has seguido los comentarios sobre la Internet de las Cosas, probablemente hayas oído al menos una vez esta asombrosa predicción: para 2020 habrá en el mundo 50.000 millones de dispositivos conectados. Hans Vestburg, exconsejero delegado de Ericsson, fue uno de los primeros en vaticinarlo en una presentación de 2010 a sus accionistas. Al año siguiente, Dave Evans, que en aquella época trabajaba para Cisco, publicó la misma predicción en un libro blanco.

La cifra actual se encuentra en algún punto entre el cálculo de Gartner, de 6.400 millones (que no incluye teléfonos inteligentes, tabletas ni ordenadores), el de la International Data Corporation, de 9.000 millones (que también excluye esos dispositivos) y el del HIS, de 17.600 millones (con todos esos dispositivos incluidos).

Algunas aplicaciones futuras de la Internet de las Cosas pueden sonar a ciencia ficción, pero entre las posibilidades más prácticas y realistas de esta tecnología se incluyen éstas:

- Avisos al móvil o a mecanismos portátiles cuando las redes de IdC detecten la proximidad de algún peligro físico.
- Automóviles que se aparquen solos.
- Encargos de alimentos u otros artículos de uso doméstico de forma automática.
- Seguimiento automático de los hábitos de ejercicio y de la actividad personal diaria, que incluya el seguimiento del progreso realizado y de las metas alcanzadas.
- Comercio de alta frecuencia entre mecanismos adaptados al efecto.

Éstos son algunos de los posibles efectos beneficiosos de la IdC en el mundo de los negocios:

- Seguimiento y localización de artículos del inventario de fábricas.
- Ahorro de combustible con el diseño de motores a gas que interactúen de modo inteligente con el medio.
- Nuevos y mejores controles de seguridad para personas que trabajen en entornos peligrosos.

La Internet de las Cosas significa que puedes coger cualquier cosa que te compres en la tienda, enchufarlo a internet e interactuar con ello. Hay tres tecnologías que hacen posible la Internet de las Cosas: los sensores, la informática y las redes.

Los sensores son cada vez más pequeños, más baratos y requieren menos energía. Los nuevos teléfonos tienen montones de nuevos sensores, y cada vez son mejores. Los sensores cuestan poco: dispositivos de todo tipo van integrando sensores ambientales, sensores de movimiento, sensores 3D... La informática va progresando a un ritmo exponencial. Hoy en día se pueden conseguir ordenadores de 10 dólares con los que jugar y experimentar, y son el equivalente a un supercomputador de los años 1990.

El problema de la IdC es la gran cantidad de productos no estandarizados que existen, que en un momento dado deberían converger. De hecho, también han aparecido nuevas redes de baja potencia y gran alcance, con sensores ideales para la IdC.

La Internet de las Cosas no va a hacer que tu nevera te hable de los alimentos que han caducado o de la leche que se te ha pasado, ni que tu lavadora hable con tu microondas. Pero puede hacer que tu cámara tome una fotografía de una agresión en la calle automáticamente y que la envíe a la policía. Otros usos pueden estar relacionados con la localización, recurriendo a pequeños sensores de pocos dólares y que se pueden fijar a cualquier cosa con enganches muy baratos (por ejemplo, a los cordones de los zapatos), o que te permitan pedir las cosas que necesitas allá donde las necesites (por ejemplo, pedir una pizza en casa a través de un imán de la cocina conectado a internet).

En el mundo industrial la IdC tiene numerosas aplicaciones. Los sensores, colocados en conductos y cisternas, efectúan controles de seguridad y de fugas, por ejemplo, para detectar al instante lo que pasa en el terreno y permitir hacer cosas que puedan cambiar la vida. La aplicación más rentable que puede tener la IdC es en la gestión de las cadenas de suplidos. Si puedes mejorar la gestión de la cadena de suplidos de una empresa, puedes ganar miles de millones.

La IdC también tendrá relación con la realidad visual y aumentada, y con las comunicaciones múltiples, que serán muy importantes para las empresas y, a largo plazo, también para los individuos.

La informática de las prendas también es parte importante de la IdC, y produce montones de datos. El Apple watch, Fitbit y otros dispositivos similares van apareciendo a gran velocidad como compañeros de vida, pero aún no nos la están cambiando. Seguimos buscando esa *app* definitiva para los mecanismos que se pueden llevar puestos. Por ejemplo, una que te diga cuándo tomarte la medicina o que controle las lámparas para que cambien de color de día o de noche y mejorar así la experiencia, o quizá incluso que te salve la vida, interactuando con los servicios de atención sanitaria. Esta tecnología IdC se está volviendo tan barata que se puede aplicar a una simple bombilla.

Los pequeños mecanismos del campo de la IdC crearán un mercado de 14 billones de dólares y producirán montones de macrodatos que habrá que analizar, procesar y a los que habrá que reaccionar para mejorar la productividad, los negocios y la vida diaria de la gente. Si los dispositivos de que disponemos tuvieran más acceso a la información sobre nuestra posición y nuestra actividad, serían mucho más útiles. Pero es necesario contar con una *app* o un servicio que funcione bien.

La Internet de las Cosas representa un concepto general que define la capacidad de los dispositivos conectados a la red de percibir y recoger datos del mundo que nos rodea, y luego compartirlos por internet, enviándolos adonde haga falta para su procesamiento y su utilización con fines diversos e interesantes.

Pero la Internet de las Cosas se basa en los datos, no en las cosas. Y éstos tienen que ser precisos, seguros y procesables. Si algo no se puede medir, no se puede gestionar. Aunque esta afirmación sea perfectamente cierta, si no se puede medir la producción de estos últimos de estos dispositivos con precisión, probablemente no tengan ningún valor, porque unos datos erróneos provocan decisiones equivocadas, especialmente cuando uno no se da cuenta de que los datos son erróneos.

Un gran ejemplo fue el accidente del vuelo 447 de Air France, en el que, según parece, los pilotos levantaron el morro del avión pese a que habían perdido velocidad, porque los sensores les dijeron que iban demasiado rápido. Este dato erróneo debió de hacer que empeoraran la situación. Contar con un equipamiento completo y conectado y generar un flujo de registros no garantiza la precisión. Los datos sólo son lo precisos que puedan ser el sistema y los sensores, combinados con cierto conocimiento del sistema, que pueda separar las señales del ruido.

Como experimento personal de la precisión de éstos, yo usé los datos de calorías de tres sensores/sistemas diferentes para hacer un seguimiento de las calorías consumidas en cinco salidas de jogging. Usando un Fitbit (con pulsómetro), otro Garmin (con pulsómetro) y uno Strava (que aprovecha el feed de datos del Garmin), obtuve tres series de datos diferentes para cada carrera. Incluso los dos que compartían datos daban resultados diferentes.

Suponiendo que hayas gestionado tus sensores y el proceso de obtención de datos correctamente para garantizar la precisión de los mismos, el paso siguiente es asegurarte de que éstos se han guardado de forma segura y de que conservan su integridad durante su tránsito por el sistema. En la IdC, los negocios pueden empezar con un sensor ya presente en las instalaciones o en el terreno, y que opere de forma remota. Del mismo modo, los vehículos pueden enviar montones de datos de telemetría y de servicio a sus fabricantes.

Todos ellos se recogen y luego pasan por una serie de servidores y aplicaciones antes de que por fin acaben en pleno centro de datos, para que sean analizados y se pueda actuar en consecuencia.

Cada minuto enviamos 204 emails, generamos 1,8 millones de «me gusta» de Facebook, enviamos 278 tuits y colgamos 200.000 fotografías en Facebook.

En 2011 se vendieron 12 millones de etiquetas RFID (usadas para capturar datos y seguir el movimiento de objetos en el mundo físico). Con el despegue de la Internet de las Cosas, se calcula que en 2021 este número habrá aumentado hasta los 209.000 millones.

El *boom* de los datos y de la Internet de las Cosas irán de la mano. Eso significa que el número de dispositivos que se conectan a la red aumentará de los 13.000 millones de hoy a los 50.000 millones en 2020.

Se espera que el sector de los macrodatos y de la Internet de las Cosas crezca de los 10.200 millones de dólares de 2013 a unos 54.300 millones en 2017.

¿Qué es lo que está pasando? La Internet de las Cosas se basa en un aumento de la comunicación entre máquinas; depende de la computación en la nube y de las redes de sensores que recogen datos; es una conexión móvil, virtual e instantánea; y dicen que va a hacer que todo sea «inteligente» en nuestras vidas, desde las farolas de las calles hasta los puertos marítimos.

En 2007 se hundió un puente en Minesota y mató a mucha gente, porque las placas de acero no estaban preparadas para soportar el peso del puente. Cuando reconstruimos los puentes podemos usar cemento inteligente: un cemento equipado con sensores para controlar las tensiones, las grietas y las torsiones, que nos avisa para que reparemos los defectos antes de que causen una catástrofe. Y este tipo de tecnología no sólo protege las estructuras del puente.

Si hay hielo en el puente esos mismos sensores del cemento lo detectarán y comunicarán la información a tu coche a través de la red inalámbrica. Si el coche sabe que se va a encontrar con un peligro más adelante, dará instrucciones al conductor para que baje la velocidad, y si el conductor no lo hace, el coche mismo lo hará. No es más que un ejemplo de las comunicaciones sensor-máquina y máquina-máquina que pueden producirse. Los sensores del puente comunican con la maquinaria del coche: convertimos información en acción.

Las implicaciones de todo esto se van haciendo evidentes. ¿Qué puedes conseguir cuando un coche inteligente y una red urbana inteligente empiezan a hablarse? Vamos a conseguir la optimización del flujo del tráfico simplemente porque en lugar de contar con semáforos con temporizadores, tendremos semáforos inteligentes capaces de responder a los cambios en el flujo del tráfico. Se enviará información sobre el tráfico y las condiciones de la red viaria a los conductores, proponiéndoles rutas alternativas a las zonas congestionadas, cubiertas de nieve o en obras.

Así que ahora tenemos sensores que monitorizan y hacen un seguimiento de todo tipo de datos; tenemos *apps* en la nube que los traducen en información útil y que la transmiten a máquinas en el terreno, lo que permite obtener respuestas móviles y en tiempo real. Y así los puentes se convierten en puentes inteligentes, y los coches... en coches inteligentes.

Las empresas de construcción han empezado a equipar sus silos y sus camiones con sensores que pueden controlar las existencias, como la cantidad de cemento almacenada, y transmitir esa información a través de una plataforma en la nube para acelerar las entregas y asegurar la continuación de la actividad. Y los gigantes de la industria petrolera han empezado a aplicar tecnologías móviles sensor-máquina que previenen accidentes con gran antelación gracias al análisis rápido y a la acción inmediata. Cuando los sensores detectan un problema —como una corrosión o una pérdida en una conducción—, la tecnología máquina-máquina (M2M) permite atajarlo inmediatamente.

Otro ejemplo de aplicación de la IdC en la industria petrolera son los pozos inteligentes. Se trata de pozos con mecanismos de control del flujo y sensores de profundidad, para que puedan ser monitorizados y controlados desde la superficie sin poner en riesgo la seguridad de los obreros. El pozo inteligente utiliza tecnología sísmica 4D, que monitoriza las fugas de gas, el flujo del agua, los cambios de presión y cualquier otra alteración provocada por las fluctuaciones en los movimientos sísmicos, facilitando la predicción y el control de los impactos sísmicos que pudieran causar daños significativos.

Pero seguimos pensando a pequeña escala. Hay que ir más allá de la construcción, más allá de la energía. Tenemos sensores capaces de medir fuerza, carga, momento dinámico y presión; sensores que pueden detectar la presencia de gas y sustancias químicas; sensores que oyen las vibraciones del sonido y distinguen entre diferentes señales acústicas; sensores que toman la temperatura, detectan el movimiento, la velocidad y el desplazamiento; que identifican la posición, la presencia y la proximidad. En otras palabras, tenemos la posibilidad de recoger datos prácticamente ilimitados en tiempo real.

¿Cómo le damos utilidad a todos estos datos? Echa un vistazo a tu propia casa. ¿Qué elementos puedes dotar de inteligencia? He aquí uno muy simple: una vez observé un sistema de videoconferencia que permitía al dueño de un perro hablarle, llamarle y darle de comer de forma remota con una aplicación inteligente. Piensa en grande. Una casa que vincula alarmas de humo, sistemas de seguridad y consolas de entretenimiento con tu teléfono. Una casa con sensores instalados en las tuberías que pueden detectar fugas antes de que se produzcan.

Recientemente Microsoft se asoció con una empresa de San Francisco para codificar información en ADN sintético para comprobar su potencial como nuevo medio de almacenamiento de registros.

Twist Bioscience le proporcionó a Microsoft 10 millones de filamentos de ADN con el objetivo de codificar datos digitales. En otras palabras, Microsoft está intentando descubrir cómo usar las moléculas que componen el código genético humano para codificar información digital.

Aunque aún estamos a años de distancia de la obtención de un producto comercial, las primeras pruebas han demostrado que es posible codificar y recuperar el ciento por ciento de datos digitales del ADN sintético. El uso del ADN podría permitir el almacenaje de enormes cantidades de ellos en un minúsculo componente. En Twist sostienen que un gramo de ADN podría almacenar casi un billón de gigabytes de datos.

Cada vez es más importante encontrar nuevos modos de almacenar datos, ya que las personas los generan cada vez más en sus vidas diarias, y porque estamos conectando millones de sensores IdC.

También es importante para Microsoft, que opera una de las mayores plataformas de nube públicas. Encontrar modos más eficientes de almacenar datos podría reducir

los costes, y el almacenamiento basado en ADN tiene el potencial de durar más que los medios usados actualmente.

Microsoft también colabora con gobiernos de todo el mundo para que entren en el mundo de los macrodatos y la IdC, cambiando así sus operaciones diarias. Los gobiernos usan herramientas de análisis predictivo para mejorar tremendamente la calidad de sus decisiones. Y van recogiendo y analizando grandes volúmenes de información para potenciar tanto su eficiencia como el servicio que dan a los ciudadanos.

En 2016 Microsoft también anunció que Cortana, su asistente personal controlado por voz, análogo a Siri, estaba a punto de emprender el vuelo gracias a una nueva asociación entre Microsoft y Boeing.

Microsoft y Boeing han anunciado su asociación para trasladar el conjunto de servicios digitales de aviación a la nube de Microsoft, y el resultado final es que los servidores de Microsoft van a volverse más inteligentes, según la empresa.

«Centralizar las aplicaciones de aviación digital de Boeing en la plataforma de macrodatos de Microsoft permitirá analizar una gran serie de datos procedentes de múltiples fuentes». El objetivo de Boeing era usar la plataforma inteligente Cortana para que los operadores de la aerolínea pudieran gestionar de forma más efectiva su inventario, para programar de forma más eficiente los horarios de pilotos y personal de cabina y los servicios de mantenimiento —que podrían provocar retrasos en los vuelos— a partir de los registros recopilados en la nube.

Imagínate un «programador inteligente» centralizado que sepa qué aviones necesitan mantenimiento y cuándo, así como la disponibilidad del personal, los horarios de los vuelos y la carga de los aviones: podría ser mucho más proactivo que la mayoría de sistemas actuales gestionados por humanos.

Quizá vayas observando una tendencia común. Una de las grandes ventajas de las tecnologías inteligentes es la capacidad de predecir y prevenir problemas cualquiera que sea su procedencia. Si la batería de tu coche fuera inteligente, te diría cuántos viajes exactamente puedes hacer antes de que se agote. Si pudieras monitorizar y controlar tus propiedades a distancia y a tiempo real, tendrías conocimiento de cualquier problema en el momento en que surgiera, y ahorrarías mucho dinero.

Este repaso no incluye ni de lejos todos los aspectos en los que van a afectarnos esta tecnología y sus datos. Va a cambiar las cosas y crear nuevas oportunidades en todos los campos imaginables.

Seguridad para todos

La Internet de las Cosas inmediatamente desata controversias sobre la privacidad de los datos personales. Se trate de información sobre nuestra ubicación física o de nuestro peso o tensión arterial —informaciones útiles quizá para quienes nos

proporcionan atención médica—, el hecho de que haya información personal nuestra actualizada y detallada viajando por el mundo es un motivo evidente de preocupación.

Si te conectaste a la red en Estados Unidos en octubre de 2016, probablemente observarías que unos cuantos de tus sitios web favoritos dejaron de funcionar durante un buen rato. Ahora los expertos dicen que se debió a que miles de dispositivos IdC —como grabadoras de vídeo o cámaras conectadas a la web— fueron víctimas de un ataque informático.

Una vez se hicieron con el control de estos dispositivos, los hackers los manipularon para que enviaran un número desorbitado de peticiones a una empresa que da servicio a los sitios web de Netflix, Google, Spotify y Twitter.

Cuando el tráfico aumentó tanto que resultó imposible gestionarlo, los sitios web se colapsaron. Fue un ataque de la vieja escuela —llamado habitualmente «ataque de denegación de servicio», pero esta vez desplegado a través de la nueva red de dispositivos que llamamos la Internet de las Cosas.

Los expertos en seguridad llevaban años advirtiéndonos de que los dispositivos conectados a internet son susceptibles de sufrir el ataque de los hackers. Lo que no sabían exactamente era qué podrían hacer los hackers una vez penetren en tu televisor, tu nevera o tu termómetro inteligentes, por ejemplo.

Ahora ya tenemos la respuesta, y es peor de lo que imaginaban los expertos. Utilizando las cámaras de seguridad y las grabaciones de vídeo de negocios de fuera de Estados Unidos, los hackers han creado un ejército de dispositivos capaces de bloquear grandes pedazos de internet y, posiblemente, silenciar los datos de las personas y de la IdC.

El condado de Montgomery (Maryland, Estados Unidos) está empleando la Internet de las Cosas para crear un campo de pruebas con el fin de ayudar a los granjeros a aumentar su eficiencia y su rentabilidad. En particular, el condado de Montgomery está trabajando con granjeros que acceden voluntariamente a implantar en sus granjas sensores que pueden ayudarles a operar de un modo más eficiente, midiéndolo todo, desde la temperatura del suelo al agua o el uso de los pesticidas. También están usando la tecnología para ayudar a los granjeros que producen lácteos a enviar los informes obligatorios para la producción de la leche.

En la otra punta del mundo, la Compañía de Aguas de Áqaba, en Jordania, que distribuye agua a 130.000 habitantes, necesitaba un sistema más efectivo de monitorización y gestión de los suministros de agua, con el fin de poder servir agua a sus clientes de un modo más eficiente. Usando la plataforma de la nube, la Compañía de Aguas de Áqaba ha encontrado un modo sencillo de hacer un seguimiento de sus infraestructuras y de sus bombas a distancia. «Ahora, con nuestra red, a través de una serie de alertas y alarmas automatizadas, podemos detectar y resolver problemas en

cuanto ocurren —explicaba Naem Saleh, director general de la Compañía de Aguas de Áqaba—. Eso significa que nos enteramos inmediatamente de cualquier problema y de su localización. No necesitamos enviar empleados en busca de fugas, lo que nos ayuda a reducir la inversión en este concepto».

Pero para poder confiar en que estas infraestructuras tan importantes puedan gestionarse y controlarse a distancia a través de la Internet de las Cosas, usando los macrodatos, es necesario asegurarse de que los datos están bien protegidos.

El gran punto débil de la Internet de las Cosas es que el número de dispositivos se multiplica. Hace diez años, la mayoría de nosotros sólo teníamos que preocuparnos de proteger nuestros ordenadores. Hace cinco, tuvimos que preocuparnos también de proteger nuestros teléfonos inteligentes. Ahora tenemos que preocuparnos de proteger el coche, los electrodomésticos, nuestras prendas y muchos otros dispositivos IdC.

El hecho de que haya tantos dispositivos susceptibles de ser pirateados significa más oportunidades para los hackers. Puede que hayas oído hablar de que los hackers podrían llegar a controlar coches de forma remotamente, haciéndolos acelerar o frenar. Pero también podrían usar dispositivos aparentemente banales como los sistemas de vigilancia de bebés o el termostato, para obtener información privada o simplemente por aguarle el día a alguien. Se trata de pensar qué podría hacer un hacker con un dispositivo si consiguiera burlar su sistema de seguridad.

Con la difusión de la Internet de las Cosas tenemos que pensar en proteger cada vez más dispositivos. Pero aunque empezaras a tomarte en serio la seguridad, las compañías tecnológicas que fabrican estos nuevos mecanismos se toman los riesgos demasiado a la ligera. Y uno de los problemas es que las empresas no actualizan sus dispositivos lo suficiente (o quizá no los actualicen en absoluto). Eso significa que un dispositivo IdC y sus datos, que eran seguros cuando los compraste, pueden volverse inseguros a medida que los hackers van encontrándoles nuevas vulnerabilidades. Los ordenadores solían tener este problema, pero las actualizaciones automáticas han contribuido a aliviar este problema.

De hecho, varios incidentes han demostrado que la Internet de las Cosas presenta numerosas vulnerabilidades, por lo que corren el riesgo de sufrir ciberataques. Los dispositivos inteligentes —electrodomésticos, coches o prendas, por ejemplo— son básicamente nuestros compañeros de vida. A diferencia de un *smartphone*, que contiene información sobre nuestras comunicaciones, contactos, fotografías y vídeos, estos otros dispositivos revelan información más específica sobre nuestra conducta, como nuestros hábitos al volante, de salud o en la cocina, o el patrón de aprendizaje de nuestros hijos. Eso ofrece unas posibilidades inmejorables para los hackers que quieran buscar datos personales, incluidos los recogidos por los dispositivos conectados a internet que se pueden llevar en las prendas de vestir.

Charlie Miller y Chris Valasek, investigadores de seguridad, alteraron para siempre el concepto de «seguridad del vehículo», cuando publicaron en Wired la demostración de que podían piratear un Jeep Cherokee de 2014 a distancia, desactivando la transmisión y los frenos. Los coches pirateados no son los únicos dispositivos de la Internet de las Cosas capaces de matar a alguien, por supuesto. El equipo médico más especializado también presenta vulnerabilidades de software y hardware que podría permitir que algún individuo malintencionado se infiltrara y los controlara, lo que podría tener consecuencias mortales. Sólo hay que preguntarle al cardiólogo de Dick Cheney que, temiéndose que el exvicepresidente pudiera ser objeto de un ataque mortal a través de su marcapasos, le desactivó la conectividad wifi durante el tiempo que Cheney ocupó el cargo.

Cuando Mattel incorporó conectividad wifi a su Hello Barbie para permitir lo que describía como conversaciones artificialmente inteligentes en tiempo real, dejó desprotegida la conexión de la *app* para *smartphone* Hello Barbie, lo que permitiría interceptar todo el audio que registra la muñeca. Una «nevera inteligente» de Samsung, diseñada para sincronizarse por wifi con el Google Calendar del usuario, no validaba los certificados de seguridad, lo que exponía al robo de las credenciales de Gmail del usuario. Incluso los dispositivos para la vigilancia de bebés siguen siendo preocupantemente inseguros, pese al riesgo de que algún pirata pueda espiar a los pequeños: un estudio de la empresa de seguridad Rapid 7 puso a prueba nueve vigilabebés y observó que todos eran relativamente fáciles de piratear.

Hace diez años yo mismo formé parte de un equipo que rediseñó el sistema de actualizaciones de Microsoft Windows para aumentar la seguridad de los datos y de los ordenadores de todo el mundo usando tecnologías P2P. Dado que Microsoft me había contratado para mejorar la ciberseguridad de millones de ordenadores, teníamos que recopilar grandes cantidades de datos y desarrollar nuevos algoritmos para comprender si los ordenadores eran capaces de ayudarse entre sí o no, qué ordenadores había que parchear primero, en qué orden habría que enviar las actualizaciones, etc. La tecnología P2P (de red de pares) había sido usada para distribuir archivos de música y películas entre consumidores, pero nunca para proteger los ordenadores de ataques.

A pesar de los tremendos esfuerzos de ingeniería realizados para proteger los datos y los ordenadores, los usuarios aún sufrían retrasos en la respuesta durante las grandes actualizaciones, que los dejaban en situación de vulnerabilidad a los ciberataques, lo que dejaba expuestos sus datos. Por ejemplo, el 13 de abril de 2004 se lanzaron cuatro parches de Windows para resolver una vulnerabilidad de seguridad que afectaba a los ordenadores que usaban Windows. El tráfico del sitio de actualizaciones de Microsoft Windows aumentó más que en la actualización anterior. Unos 3 o 4 millones de usuarios visitaron el sitio. Como resultado de esta sobrecarga,

muchos clientes se encontraron con que su perfil de Windows Update no funcionaba bien y no pudieron descargar los últimos parches, lo que dejaba sus ordenadores y sus datos en una situación de riesgo frente a los hackers.

El retraso que sufrieron los clientes implicó que algunos tuvieran que esperar horas o incluso días antes de instalar el parche en su sistema, alargando así el tiempo de exposición de estos sistemas vulnerables más de lo deseado y abriendo la puerta a los piratas informáticos. Las actualizaciones de seguridad, el software antivirus y otras actualizaciones críticas tenían que llegar a los ordenadores lo antes posible para minimizar el riesgo de propagación de gusanos informáticos, de robo de datos o de ciberataques.

Hasta el año 2000, las empresas de seguridad que distribuían soluciones a los ordenadores de todo el mundo lo hacían situando equipos específicos en determinados lugares de la red, o en sus límites, el mejor ejemplo de este tipo de soluciones es, de nuevo, Akamai, que gestiona más de 40.000 servidores en más de 40 países. No obstante, usar redes de pares y hacer que los ordenadores colaboren protegiéndose mutuamente es en sí mismo un sistema expansible y mucho más rápido. Internet puede adaptarse espontáneamente a la demanda aprovechando los recursos que proporciona cada ordenador, y aumentando su nivel de protección mucho más rápidamente. La capacidad de mantener segura la red crece al mismo tiempo que la demanda de esas soluciones de seguridad, ya que los ordenadores se ayudan mutuamente al tiempo que se protegen.

La última versión de Microsoft Windows ha llenado numerosos titulares últimamente, pero uno de ellos reveló un modo totalmente novedoso de hacer llegar actualizaciones de Windows al usuario: la comunicación entre pares (*peer-to-peer* o P2P). En el pasado, Microsoft ha invertido mucho en sus propios servidores de actualización, que proporcionan parches y otros remedios a los usuarios, sean particulares o empresas. Pero ahora mismo está usando la tecnología que diseñamos hace más de una década para proteger los datos más rápidamente.

El próximo desafío de la Internet de las Cosas es protegerlo todo de los gusanos, los virus y los ataques que circulan por la red. Para conseguirlo, empresas como Microsoft tienen un mecanismo activo a todas horas, 365 días al año, para monitorizar las nuevas vulnerabilidades informáticas, el *malware* y los defectos potenciales que puedan convertir a tu ordenador en víctima de un ataque informático. Una vez detectada la vulnerabilidad, el sistema de actualización de Windows se pone en marcha. Este sistema de protección cibernética es el mayor del mundo y se ocupa de la seguridad de 400 millones de ordenadores, pero ahora tendrá que ampliarse a la Internet de las Cosas, para mantener el alma de sus datos protegida y segura.

3

Inteligencia artificial: humanos contra máquinas

Tiene cuidado con lo que lee, pues eso es lo que escribirá. Tiene cuidado con lo que aprende, porque eso es lo que sabrá.

ANNIE DILLARD

IA

Yo pertenezco a una generación que ha sido educada y que ha crecido profesionalmente con una estructura lineal, paso a paso, haciendo primero una cosa muy bien y dominándola antes de pasar a la siguiente. Así, primero dediqué cinco años a estudiar para obtener la licenciatura como ingeniero de telecomunicaciones, luego me especialicé durante cuatro años más doctorándome en redes e informática, y luego aprendí sobre emprendeduría en California, trabajando en diversas empresas emergentes y descubriendo cómo buscar financiación, como presentar ideas o como vender empresas, y por último pasé casi una década gestionando y dirigiendo equipos de innovación de alto rendimiento, sacando lo mejor de la gente y de su talento, afrontando los complejos recorridos emocionales de los individuos y de las empresas a medida que crecen y se enfrentan a la incertidumbre, al éxito o al fracaso.

Últimamente he iniciado un nuevo viaje, el del autodescubrimiento, el de mirar en mi interior, el de buscar significado, sensación de pertenencia, valores, el convencimiento de estar haciendo cosas que tengan un sentido más allá del de ganar dinero o hacer que tu producto o invento llegue a millones de personas, concentrándome más en cómo generar un impacto social positivo en el mundo. Utilizando herramientas nuevas, como la autoconciencia, la conexión emocional, la empatía, conectando con los más necesitados e intentando poner todo lo aprendido durante una larga carrera llena de conocimientos al servicio de otros y devolver lo recibido, concentrándolo en hacer el bien, en lugar de buscar hacerlo bien. Y es precisamente ahora, veinte años después de iniciar mi recorrido paso a paso (educación, ciencia, emprendimiento, liderazgo, autoconciencia), cuando puedo aunarlo todo para crear un espacio de fuerza positiva en la vida de la gente.

Cuando comparo mi recorrido con el de generaciones posteriores, por ejemplo los *postmillennials*, veo diferencias radicales. Los *millennials* o *postmillennials* son un grupo de personas nacidas aproximadamente a partir de 1984 o de mediados de los años noventa, respectivamente, o después. Así pues, hay una o dos generaciones de

diferencia entre la mía (la generación X) y la de ellos. Cuando interactúo con estas generaciones más jóvenes, en mi trabajo diario, en conferencias, congresos o reuniones, enseguida me doy cuenta de que todo lo que yo he hecho en un proceso lineal, paso a paso, ellos lo hacen a la vez, y lo encarnan desde el principio, lo que hace difícil situarlos en una categoría determinada o describirlos.

Suelen ser autodidactas, algunos de ellos incluso han dejado los estudios pronto, lo cual no significa que no tengan un alto nivel educativo, sino que más bien lo han adquirido a través de cursos en internet, vídeos, libros y artículos. Son proactivos, y la ingeniería y los prototipos son algo que hacen por defecto, con la mentalidad de construir cosas, usando componentes de hardware y software baratos para conseguir lo que necesitan rápidamente. También son emprendedores sin haber tenido la necesidad de acudir a una escuela de negocios o sacarse un máster, ya que enseguida se plantean cómo lanzar una pequeña empresa, ponerla en manos de la gente y repetir el proceso rápidamente, incluso recurriendo a financiación a través de plataformas de *crowdsourcing* o pequeñas fundaciones emergentes. Les encanta la ciencia y suelen leer textos científicos o de divulgación. Desde un principio emprenden una trayectoria que les lleva a buscar conectar más consigo mismos, encontrar un objetivo y un significado, y desarrollar una buena inteligencia emocional a través de la meditación, el yoga o la concienciación plena. Intentan implicarse o dirigir iniciativas y proyectos con un impacto social o medioambiental (en países emergentes como la India u otros de Latinoamérica o África), y están convencidos de que haciendo el bien puedes llegar a hacerlo bien.

Hacer todo esto de golpe, sin una formación formal en ninguna de las disciplinas implicadas (por ejemplo, ingeniería, ciencia, iniciativa empresarial o liderazgo) puede parecer imposible. Lo que yo, y otros como yo, tardamos casi tres décadas en conseguir, ellos lo intentan hacer por defecto desde el inicio de su carrera profesional. Usan la ciencia, la ingeniería, la gestión empresarial, la autoconciencia y un enfoque social para hacer algo que tenga un impacto. ¿Cómo es posible? Una respuesta podría ser que, haciendo muchas cosas a la vez, no hacen ninguna en profundidad, por lo que los resultados son muy superficiales o poco consistentes. Pero sé por experiencia que no es ése el caso. Están tomando el timón de nuestro futuro, se están convirtiendo en los nuevos líderes del mundo, con resultados y logros muy coherentes, y arraigados en profundos valores. Son la generación que vivirá en un mundo impulsado por la inteligencia artificial.

La respuesta está en la forma en la que aprenden. A diferencia de mi generación, que intentaba entender las cosas muy bien, construir modelos complejos del mundo o de una disciplina que pudieran explicar cómo funciona un campo científico determinado, la gestión de una empresa o un modelo de ingeniería informática, ellos aprenden iterando y poniendo en práctica el principio de Pareto.

El principio de Pareto (también conocido como regla del 80/20) afirma que, en muchos casos, aproximadamente el 80 por ciento de los efectos se deben al 20 por

ciento de las causas, o que se puede obtener el 80 por ciento de unos conocimientos sólo comprendiendo el 20 por ciento de las cosas más interesantes. Joseph M. Juran, teórico de la gestión empresarial, fue quien sugirió este principio y le puso el nombre del economista italiano Vilfredo Pareto, que señaló esta conexión 80/20 en 1896 en la Universidad de Lausana, y que la publicó en su primer estudio, «Cours d'économie politique». Básicamente, Pareto demostró que aproximadamente el 80 por ciento del territorio italiano era propiedad del 20 por ciento de la población; Pareto desarrolló el principio observando que aproximadamente el 20 por ciento de las vainas de guisante de su jardín contenían el 80 por ciento de los guisantes.

Este principio se ha aplicado a muchas disciplinas. Por ejemplo, Telefónica observó que encontrando solución al 20 por ciento de los errores más recurrentes de un sistema determinado se eliminarían el 80 por ciento de los fallos relacionados. En el campo del aprendizaje, el principio de Pareto plantea que el 80 por ciento de los resultados en cualquier iniciativa son producto del 20 por ciento de los esfuerzos. Por ejemplo, al estudiar un libro, se puede aplicar el principio de Pareto o regla del 80-20: el 20 por ciento del libro corresponderá a los títulos de las secciones, las palabras claves y las imágenes. No obstante, este 20 por ciento te permitirá captar la idea principal de la materia que estás estudiando, lo que se traduce en el 80 por ciento del aprendizaje o de la comprensión que puedes alcanzar con el estudio del mismo.

Aplicando este principio puedes cambiar significativamente tu manera de aprender. Puede ayudarte a leer más rápido y a adaptarte mejor. En lugar de tener que aprender muchas habilidades diferentes a la perfección y de forma lineal, puedes abordar muchas de ellas en paralelo y extraer los fragmentos más importantes, explorar los resultados y aplicar rápidamente lo aprendido. Eso es parecido a lo que hacen los algoritmos de IA. En lugar de intentar construir modelos perfectos y complejos del mundo e intentar comprenderlos, que es lo que hacía mi generación con el aprendizaje formal, la nueva generación aprende a partir de la experiencia y los experimentos, con datos, adaptando su aprendizaje y sus modelos para mejorar el rendimiento. Es un nuevo modelo de pensamiento, un nuevo paradigma de estudio, de trabajo y de aprendizaje, y un nuevo modelo generacional que yo no había visto antes y que me fascina. Algo que yo sólo he podido hacer a mitad de mi carrera, y que las nuevas generaciones hacen por defecto desde el principio. Y, curiosamente, está ocurriendo en paralelo con la emergencia de nuevos modelos de aprendizaje para los ordenadores y las máquinas, la inteligencia artificial.

Es mediodía de un soleado día de verano. La gente está sentada en las terrazas, los coches huyen de la ciudad para aprovechar el largo fin de semana que se presenta en la costa. Yo estoy a punto de pedirme una ensalada y una agua con gas cuando llega la persona que estoy esperando. Es uno de esos *postmillennials*, un chico de veinte años que acaba de incorporarse a mi equipo. Al igual que todos los miembros de mi

equipo, intento quedar para desayunar, almorzar o cenar con los recién llegados para conocerlos y escuchar qué es lo que les apasiona y qué quieren hacer en la siguiente fase de sus vidas. Este chico es el típico ejemplo de enfoque no lineal que he descrito antes. Ha dejado los estudios, ha aprendido programación de forma autodidacta, ha lanzado una pequeña empresa en África y la ha vendido después, contribuyendo con ello a llevar agua fresca a algunos poblados remotos, y ha aprendido sobre inteligencia artificial para intentar mejorar algunos videojuegos.

Me cuenta que un día, mientras jugaba a su videojuego favorito, se dio cuenta de que perdía todo el rato. Y no importaba lo que llegara a dominar el juego; había algunos personajes del juego que parecían invencibles, ya que iban mejorando sus habilidades a medida que progresaba el juego, hasta acabar siendo difícilísimos de batir. Consultando diversos foros de internet, se dio cuenta de que detrás de esos personajes de videojuego no había humanos, como él, sino robots de inteligencia artificial preprogramados que podían jugar en nombre de otro jugador, incluso cuando el jugador no estaba presente.

Y esos robots programados del juego podían aprender y adaptarse; básicamente usaban sencillas versiones de algoritmos de inteligencia artificial que leían los datos del juego y rápidamente aprendían nuevos trucos para ganar el juego. Así que decidió crearse su personaje basado en la IA, para aumentar las posibilidades de ganar en el videojuego. Invirtió meses, y al final alcanzó una de las puntuaciones más altas de la historia del videojuego, reveló sus descubrimientos a la comunidad de desarrolladores y jugadores y compartió su experiencia con otros. Y, sin saberlo, por el camino, se convirtió en un verdadero experto en inteligencia artificial: comprendía los conceptos científicos que había tras esos algoritmos, estudió artículos científicos sobre el funcionamiento cerebral para hacer que aquellos algoritmos artificiales jugaran aún mejor, leyó sobre el cerebro y se embarcó en un viaje de constante aprendizaje y experimentación.

Yo tenía curiosidad por saber si había usado alguna metodología particular para aprender. Y él me dijo que se había concentrado en seguir a algunas personas que habían tenido una trayectoria como la mía e interactuar con ellos, profesionales rigurosos, que habían ido creciendo en diversas disciplinas paso a paso y que habían sacado lo mejor de cada una, que había aplicado la regla de Pareto (80/20) y que había repetido el proceso una y otra vez para seguir aprendiendo. Ésta es una generación que personifica la inteligencia artificial y el aprendizaje de la máquina por defecto, aplicando e interactuando con una nueva forma de inteligencia, una tecnología y una disciplina que condicionará a su generación y a muchas otras generaciones futuras.

De hecho, todo el día, desde el momento en que te despiertas al momento en que te duermes, estás rodeado de IA. Tu despertador da la alarma por la mañana. Hace sonar una canción que no has oído nunca, pero resulta que te gusta. Cortesía de Spotify, que ha estado tomando nota de tus gustos musicales, como si fuera tu locutor

de radio personal. Lo más probable es que la canción también haya sido creada con ayuda de la IA, como ocurre con muchas de las películas de Netflix, que adaptan el guión según lo que gusta o no gusta a la gente. Mientras desayunas y lees el periódico de la mañana, ves unas noticias que han sido seleccionadas usando un algoritmo de IA. La temperatura de tu casa es perfecta, y tu factura eléctrica se ha reducido notablemente desde que has instalado un termostato Nest, que usa algoritmos para saber cuándo estás en casa, y cuándo poner en marcha la calefacción para que la casa esté a punto cuando llegues, sin que ello suponga hinchar el gasto energético.

Mientras conduces hasta el trabajo, tu coche ajusta constantemente la ruta usando Waze, un sistema de predicción de tráfico, para reducir al mínimo el viaje en hora punta. En el trabajo, la IA te ayuda a filtrar los documentos más importantes y clasificarlos según el contexto del día en su web interna. El buzón de entrada de tu correo electrónico resume convenientemente los mensajes más importantes para ti, dejando el resto en otras carpetas. Tu navegador web te sugiere automáticamente palabras clave de búsqueda en el momento en que escribes la que quieres buscar, y te ayuda a completar tu búsqueda. Eso, a su vez, te ayuda a encontrar un sitio web que te ahorra dinero en tu próximo viaje de vacaciones. Y sin darte cuenta le has sacado mucho más partido a tu tiempo, hora por hora, de lo que habrías hecho sin ayuda de esos algoritmos de IA.

En este mismo momento, los algoritmos de IA están trabajando duro, intentando predecir tu próximo movimiento, aprendiendo a conducir y luchando contra el cáncer. Allá donde haya una base de macrodatos, probablemente haya un algoritmo de inteligencia artificial trabajando duro para aprender lo que pueda de esos datos. El aprendizaje de las máquinas, parte integrante de la inteligencia artificial, es un intento por programar los ordenadores para que aprendan de los datos. El aprendizaje de las máquinas es la nueva infraestructura que hay detrás de todo. La geopolítica y el cambio climático afectarán al mercado laboral, pero no tanto como afectará el aprendizaje de las máquinas al panorama económico.

El aprendizaje de las máquinas es la capacidad de éstas para aprender, en lugar de decirles qué hacer, y una de las cosas que aprenden a hacer muy bien es detectar tendencias y patrones a partir de los datos. Estos patrones, si se basan en series de datos lo suficientemente largas y en datos de suficiente calidad, pueden usarse para predecir el futuro. Sí, el aprendizaje automático mira al pasado para encontrar patrones e intentar predecir el futuro. No siempre funciona, pero sí en muchos casos. Los avances en los algoritmos, la computación y la enorme cantidad de registros transforma industrias enteras.

Y la IA presenta cada vez más posibilidades de aumentar su impacto, con el crecimiento exponencial de la potencia informática de la que disponemos. De hecho, los ordenadores de 1.000 dólares de hoy en día superan ya el cerebro de un ratón, y están a una milésima parte del nivel de los humanos. Eso no parece mucho, hasta que recordamos que estábamos a una billonésima parte del nivel humano en 1985, a una

milmillonésima en 1995 y a una millonésima en 2005. Estar a una milésima en 2005 hace pensar que para 2025 podremos disponer de un ordenador asequible que rivalice con el poder del cerebro humano.

Si eres como yo, antes seguramente pensabas que la inteligencia artificial era un concepto de ciencia ficción, pero últimamente se está convirtiendo en un término más popular que tiene confundida a mucha gente. Hay tres motivos por los que el término IA confunde a tanta gente.

1. Asociamos la IA con las películas. *La Guerra de las galaxias*. *Terminator*. *2001: Una odisea del espacio*. Y eso es ficción, al igual que los personajes de los robots. De modo que hace que la IA nos resulte algo ficticia.
2. La IA es un concepto muy amplio. Va desde los algoritmos que usa la calculadora de tu teléfono y los coches que se conducen solos a algo que quizá en el futuro cambie el mundo espectacularmente. La IA hace referencia a todas esas cosas, lo cual resulta confuso.
3. Usamos constantemente la IA en nuestra vida diaria, pero en muchos casos no nos damos cuenta de que es IA John McCarthy, que acuñó el término «inteligencia artificial» en 1956, se quejaba de que «en cuanto funciona, la gente deja de llamarla IA».

Tenemos que dejar de pensar en robots. Un robot es un contenedor para la IA, que a veces imita la forma humana —a veces no—, pero la IA en realidad es el ordenador que está en el interior del robot. La IA es el cerebro, y el robot es el cuerpo, si es que tiene cuerpo. Por ejemplo, el software y los datos que hay detrás de Siri son IA, la voz de mujer que oímos es la personificación de esa IA, y no hay ningún robot.

Un reciente estudio realizado con Google Trends muestra un curioso aumento del nivel de interés en el aprendizaje de las máquinas y la IA con el paso del tiempo. Se puso de moda hacia 2005 y luego hubo un período de enfriamiento, pero en cuanto los macrodatos empezaron a ganar notoriedad, hacia 2010, se creó una tendencia creciente que sigue hasta la fecha. Lo bueno es que el «aprendizaje de las máquinas» en realidad no es más que la confluencia de diversas disciplinas relacionadas como la informática, la estadística y la matemática.

Una pregunta que surge espontáneamente es por qué experimentamos un crecimiento tan explosivo de aprendizaje automático y de sus aplicaciones actualmente, pese a que la tecnología lleva vigente entre treinta y cuarenta años. Algunos de los impulsos que aceleran el aprendizaje automático con la consecuencia no buscada del cambio del escritorio a la web, es decir, que actualmente acumulamos de forma automática grandes cantidades de datos prácticamente sobre cualquier cosa. Puede parecer algo evidente, pero no siempre valoramos el impacto que supone en toda su dimensión.

Si en 1990 hubieras sido un brillante investigador sobre el aprendizaje automático y hubieras trabajado en el Microsoft Word, los únicos datos de que dispondrías serían

los que podías recoger en el laboratorio. Si descubrías una gran mejora, tardarías años en lanzarla. Diez años más tarde, el mismo investigador tiene acceso a una amplia colección de consultas, clics, visionados de página, sitios web y vínculos, y si esa persona puede encontrar una mejora algorítmica para potenciar el rendimiento de una operación determinada usando los macrodatos, puede desarrollarla al instante.

Las compañías tecnológicas desplegadas en la red fueron las primeras en aplicar la tecnología del aprendizaje automático a enormes cantidades de datos, motivo por el que aparecieron tecnologías como MapReduce y BigTable, creadas por Google, que procesan datos clave, aunque estamos viendo las mismas técnicas en otros sectores, como la fabricación, las finanzas, la venta al detalle, la energía, la administración, la asistencia sanitaria, la biología, la seguridad y muchos otros.

Además, hoy en día la computación es barata y está al alcance de todos. Los algoritmos y los métodos usados para sacar partido a esos datos no han cambiado en esencia desde que empezaron a usarse, en 1974, pero actualmente tenemos una potencia de CPU un millón de veces mayor. Muchos de los algoritmos avanzados que impulsan el aprendizaje automático y la inteligencia artificial en el mundo son ahora de código abierto y por tanto gratuitos para todos. Por ejemplo, hace poco un grupo encabezado por Elon Musk, fundador de Tesla Motors, y Sam Altman, presidente de Y Combinator, reveló la existencia de una organización sin ánimo de lucro con un valor de 1.000 millones de dólares llamada OpenAI que ha prometido compartir todos sus estudios y tecnología sobre IA con el público. Por su parte, Google y Facebook han abierto al público muchos de los algoritmos que usan sus sistemas de aprendizaje automático.

Los investigadores ya estaban emocionados con el aprendizaje automático en los años cincuenta, y ya en aquel entonces estaban claras muchas de sus implicaciones. Pero para alcanzar ese logro hacía falta una innovación tecnológica continuada de la que formaban parte el microprocesador, el ordenador personal, el tener un ordenador sobre cada escritorio, la computación en red, el acceso generalizado a internet, el explorador web, el comercio electrónico, los motores de búsqueda y las redes sociales.

En última instancia, el motivo por el que se ha alcanzando este nivel en el aprendizaje automático y la inteligencia artificial no es ningún avance específico en los algoritmos, sino más bien la amalgama de tecnologías que permiten la digitalización y la adaptación a la red de los datos, lo cual está empezando a dar sus frutos.

Algoritmos

Los algoritmos gobiernan nuestras vidas cada vez más. Nos encuentran libros, películas, trabajos y citas, gestionan nuestras inversiones y descubren nuevos

fármacos. Cada vez más, estos algoritmos aprenden de los rastros de datos que vamos dejando en nuestro nuevo mundo digital, aprenden del alma de éstos. Como niños curiosos, nos observan, imitan y experimentan. Y en los laboratorios de investigación y las universidades más importantes del mundo se libra una carrera por inventar algoritmos nuevos y mejores, capaces de descubrir cualquier tipo de conocimiento a partir de los datos, y de hacer cualquier cosa que deseemos antes de que la pidamos siquiera.

Una receta de cocina es un algoritmo para humanos. No obstante, los ordenadores y las máquinas necesitan algoritmos mejor definidos que les ayuden a completar las tareas tal como se han definido, y que siempre den el mismo resultado partiendo de los mismos parámetros. En este sentido, una receta de cocina no es un buen algoritmo para un ordenador, ya que el ordenador debería ser capaz de identificar los ingredientes a través del reconocimiento visual o táctil, algo que no queda definido en una receta; tendría que saber cómo mover los ingredientes de un lado al otro, etc.

Hoy en día usamos algoritmos en nuestra vida diaria de modos que nos resultan inimaginables. Pongamos que buscas un apartamento en San Francisco. El auge del sector tecnológico y las estrictas leyes urbanísticas que limitan las nuevas construcciones hacen que la ciudad sea cada vez más cara y, por lo que dice la mayoría, cada vez es más difícil encontrar ofertas interesantes. Los anuncios aparecen y desaparecen rápidamente, a las jornadas de puertas abiertas para mostrar un piso o una casa acuden multitud de personas, y en muchos casos las llaves acaban en manos del primero que consigue depositar un talón en manos del vendedor.

Este mercado tan salvaje deja poco espacio para la comprobación de datos y la deliberación que supuestamente debería caracterizar la actuación de un comprador racional. A diferencia de un cliente de un centro comercial o un comprador por internet, que puede comparar opciones antes de tomar una decisión, el aspirante a vecino de San Francisco tiene que decidirse rápidamente: puede quedarse el apartamento que está viendo, abandonando la posibilidad de considerar todos los demás, o puede salir por la puerta y no habrá una segunda oportunidad. Y para hacer eso, necesita un algoritmo.

Para simplificar las cosas, supongamos por un momento que sólo nos preocupa ampliar al máximo las probabilidades de conseguir el mejor apartamento disponible. ¿Cómo vamos a saber si un apartamento es realmente el mejor si no tenemos una referencia? ¿Y cómo vamos a establecer esa referencia si no examinamos (y dejamos pasar) unos cuantos apartamentos? Cuanta más información reúnes, mejor reconocerás la ocasión ideal cuando la veas, pero más probable será, al mismo tiempo, que ya la hayas dejado pasar.

Cuando se les presenta este tipo de problema, la mayoría de personas, intuitivamente, dicen algo así como que hace falta encontrar un equilibrio entre observar y lanzarse a por él: que tienes que ver los suficientes pisos como para crearte una referencia, y luego quedarte con el primero que se ajuste a la referencia que te

has creado. De hecho esta noción de equilibrio es muy correcta. Lo que la mayoría no consigue determinar con precisión es dónde está ese equilibrio. Afortunadamente, hay un algoritmo que tiene la respuesta.

Si quieres tener las máximas probabilidades de conseguir el mejor apartamento posible, dedica el 37 por ciento del tiempo de búsqueda (once días, si te has dado un mes para la búsqueda) explorando posibilidades sin comprometerte. Deja el talonario en casa; sólo estás calibrando. Pero a partir de ese momento, prepárate para dejar una paga y señal inmediatamente en el primer lugar que veas que sea mejor que lo que has visto anteriormente. No se trata de alcanzar un compromiso entre la observación y la decisión final que te satisfaga de forma intuitiva. Es la solución óptima demostrada por Brian Christian y Tom Griffiths en *Algorithms to Live By*.

Y resulta que la búsqueda de vivienda es uno de esos problemas de nuestra vida diaria que requieren un algoritmo. Pero hay otros. ¿Cuántas veces vale la pena dar la vuelta a la manzana antes de quedarse con un aparcamiento? ¿Hasta dónde conviene forzar la suerte en un negocio arriesgado antes de hacer caja y retirarse? ¿Cuánto tiempo conviene esperar para obtener una oferta mejor por la casa o el coche? El mismo desafío también se presenta en un ambiente más peliagudo: las citas. Poner fin a las relaciones en el momento justo es toda una ciencia, el secreto de la monogamia en serie.

La gente se enfrenta a estos problemas a diario, y en algunos casos les resultan angustiosos. Pero la angustia es innecesaria. Matemáticamente, al menos, son problemas que tienen solución.

Cada agente inmobiliario, cada conductor y cada abogado que ves a tu alrededor utiliza un algoritmo para desempeñar su trabajo diario. Existe una serie de problemas a los que se enfrenta todo el mundo, problemas que son el resultado directo del hecho de que nuestras vidas transcurren en un espacio y un tiempo finitos. ¿Qué deberíamos hacer, y qué deberíamos dejar por hacer, en un día o en una década? ¿Hasta qué punto deberíamos aceptar el caos y cuándo empieza a ser excesivo el orden? ¿Dónde está el equilibrio perfecto entre probar nuevas experiencias y repetir las que más nos gustan para sentir que tenemos una vida completa?

Puede parecer que son problemas exclusivos de los humanos, pero no lo son. Durante más de medio siglo, los ingenieros informáticos han estado afrontando —y resolviendo en muchos casos— los equivalentes de estos dilemas de la vida diaria. ¿Cómo puede fijar la «atención» un procesador para ejecutar todas las tareas que le pide el usuario, con el mínimo trabajo y en el menor tiempo posible? ¿Cuándo debería pasar de una tarea a la otra y cuántas tareas debería abordar a la vez en primera instancia? ¿Cómo puede aprovechar al máximo sus limitados recursos de memoria? ¿Debería recopilar más datos, o actuar basándose en los datos que ya tiene? Aprovechar el tiempo puede ser un gran desafío para los humanos, pero los ordenadores que tenemos a nuestro alrededor aprovechan cada milisegundo sin

pensárselo dos veces usando algoritmos de inteligencia artificial muy potentes. Y podemos aprender mucho viendo cómo lo hacen.

Aprendizaje automático

Un subconjunto particular de todos los algoritmos es el de los usados para el aprendizaje automático. Los algoritmos que usan las máquinas para aprender son la automatización del descubrimiento —el método científico hinchado de esteroides— que permite que los robots inteligentes y los ordenadores se programen a sí mismos. Ningún campo científico actual tiene tanta importancia ni está rodeado de tanto misterio. Hoy en día estas sofisticadas máquinas con capacidad de aprender son las responsables del funcionamiento de Google, Amazon y de tu teléfono móvil. El aprendizaje automático es lo que hace que nuestros teléfonos tengan tantas prestaciones, que Netflix te sugiera películas y que salgan elegidos los presidentes. Pero es necesario que usemos también el aprendizaje automático para hacer aún más: para curar el cáncer y el sida y, posiblemente, para solucionar todos los problemas que tiene la humanidad.

Pedro Domingos es profesor de Aprendizaje Automático en la Universidad de Washington y esto lo entiende muy bien. Él busca el algoritmo maestro, un algoritmo que en principio pueda usarse para descubrir conocimientos a partir de datos en cualquier dominio. Él recuerda que quizá no lo sepamos, pero el aprendizaje automático está por todas partes. Cuando tecleas unas palabras en un motor de búsqueda, el motor se basa en eso para deducir qué resultados darte (y qué anuncios, además). Cuando lees el correo electrónico no ves la mayor parte del spam, porque el aprendizaje automático ha hecho posible que el correo llegue ya filtrado. Ve a Amazon.com a comprar un libro o a Netflix para ver un vídeo, y un algoritmo de aprendizaje automático te recomendará algunos que quizá te gusten. Facebook usa el aprendizaje automático para decidir qué actualizaciones mostrarte, y Twitter hace lo mismo con los tuits.

Tradicionalmente, el único modo de hacer que un ordenador haga algo es comprender muy bien cómo funcionan las cosas y escribir un algoritmo que explique cómo funcionan. Para eso es necesaria una comprensión profunda de los principios que rigen un sistema determinado, por ejemplo, su mecánica, sus interacciones o sus reglas de funcionamiento. No obstante, hay cosas que son tan complejas que no resultan fáciles de comprender, y menos aún de explicar. Los algoritmos de aprendizaje automático trabajan de un modo diferente: en lugar de intentar explicar por qué funcionan las cosas y presentar una ecuación que las describa, lo descubren

por su cuenta haciendo deducciones a partir de los datos. Y cuantos más tienen, mejores son. ¡En cierto sentido, se programan a sí mismos!

El aprendizaje automático es una tecnología que se construye a sí misma. Desde la antigüedad, los humanos han sido los únicos que han diseñado artefactos. Pero los algoritmos de aprendizaje automático son capaces de diseñar otros artefactos. Cuantos más datos tienen, mejor es el resultado de lo que aprenden.

En esencia, el aprendizaje automático es pura predicción: predice lo que necesitamos, nuestras próximas vacaciones, las habilidades que nos serán útiles a medida que cambie el mundo. El aprendizaje automático hace todo esto basándose en los datos y las experiencias del pasado. Cuanto más registros tiene, cuanto mejores sean, más probabilidades tendrá de predecir el futuro con una base científica fiable.

No obstante, estas predicciones también tienen limitaciones relativas a los datos que pueden recogerse. Los macrodatos y el aprendizaje automático amplían en gran medida las posibilidades ampliando la dimensión de los datos disponibles para predecir cosas nuevas en un futuro. Algunas de ellas podríamos predecirlas mentalmente sin gran esfuerzo, desde la captura de una bola hasta las sencillas previsiones económicas en nuestra vida diaria. Otras, en cambio, son impredecibles, porque son aleatorias. Por ejemplo, el número que saldrá al tirar un dado. Pero hay un gran número de cosas que quedan en el espacio intermedio, y es ahí donde pueden ser de gran ayuda el aprendizaje automático y la IA.

La idea básica que hay tras los macrodatos y la IA es que hoy en día consumidores y empresas dejamos constantemente un rastro digital, un «alma de datos» de la que pueden sacarse elaboradas conclusiones. Aunque la idea es de gran calado, la parte más importante del proceso a menudo pasa desapercibida. Se trata del proceso de convertir los obtenidos en información útil.

Los macrodatos no se convierten solos en información útil. Los macrodatos no crean beneficios directos. El aprendizaje automático y los algoritmos de la IA pueden hacer buenas predicciones, sí, pero es igualmente importante hacer las preguntas correctas. Una cosa es saber jugar con los números, pero más importante aún es comprender qué conclusiones se pueden obtener de estos números que sean útiles para el negocio, y qué acciones se pueden tomar a partir de esas conclusiones. Es importante tener experiencia y un buen conocimiento del negocio, pero también curiosidad y pasión.

Aprendizaje cognitivo

El número de transistores de un ordenador cada vez se acerca más al número de neuronas en un cerebro humano, aunque el cerebro gana en cuanto al número de conexiones. En un microprocesador, cada transistor suele estar conectado a unos pocos más. En cambio, una neurona tiene miles de sinapsis que la conectan con otras

neuronas. Los cerebros son realmente rápidos revisando opciones, encontrando coincidencias y adaptándose a nuevos contextos.

Eso no significa que no podamos intentar hacer paralelismos entre un cerebro y un ordenador. De hecho, hoy en día una de las principales limitaciones de los ordenadores en comparación con los cerebros es la energía. Tu cerebro usa mucha menos energía de la que usa un ordenador para realizar la misma tarea.

Los ordenadores, gracias al aprendizaje automático, hacen predicciones basadas en datos, pero también las hace el cerebro. De hecho, el cerebro hace predicciones constantemente, por necesidad. Llegas a la parada del autobús y ves que alguien lleva esperando tres minutos. ¿Cuánto tardará en llegar el próximo autobús teniendo en cuenta lo que lleva esperando otro? Vivimos en la era de los macrodatos, en la que los ordenadores encuentran patrones invisibles a simple vista. Pero nuestro día a día está lleno de «microdatos». El cerebro necesita tomar decisiones muchas veces a partir de informaciones mínimas, en muchos casos sólo con un dato. ¿Cómo lo hacemos?

La posibilidad de hacer predicciones a partir de pequeños datos se ve, por ejemplo, en el juego. Por ejemplo, si compramos un solo billete en una lotería, ¿Qué probabilidad tenemos de ganar?, ¿tenemos una probabilidad de uno contra uno de ganar? ¿Del ciento por ciento? Poder predecirlo suena erróneo y demasiado optimista. ¿Pero hasta qué punto lo es?

Intentar usar los billetes de lotería premiados y los que no lo han sido sería razonar hacia atrás, pero antes tenemos que intentar razonar hacia delante a partir de hipótesis. Es decir, necesitamos comprender cuál es la probabilidad de tener un billete ganador, cuál es la probabilidad de que eso ocurra. Por ejemplo, podemos suponer que la administración de loterías querrá dar billetes ganadores sólo al 1 por ciento de la gente, en lugar de al 90 por ciento (con lo que difícilmente ganarían dinero). Eso reduce las opciones de contingencia, y hace que el cerebro tome decisiones basándose en muy poca información.

A esto lo llamamos «probabilidad *a priori*», aunque estemos lanzando una suposición al vuelo. En realidad, resulta difícil analizar una situación para la que el cerebro no tiene ninguna referencia. Y ése es exactamente el proceso al que se enfrenta nuestro cerebro cuando intentamos predecir el futuro. Todo el conocimiento que obtenemos a partir de los datos es por definición incierto, porque de algún modo estamos tratando de inferir cosas. Lo hacemos constantemente, a través de la intuición, transfiriendo conocimiento de una situación a otra e intentando predecir lo que sucederá al rellenar esos huecos.

El cerebro también usa otras reglas y algoritmos para hacer predicciones. Una de ellas se basa en el principio de Copérnico. Copérnico sostenía que si te encuentras con un evento al azar, lo más probable es que haya aparecido en tu vida cuando ya ha alcanzado la mitad de su tiempo de ocurrencia, por lo que le quedará de vida el tiempo que haya durado ya. Esta regla puede ayudar a predecir cosas. Por ejemplo, significaría que Google podría durar hasta 2032, y que la relación con ese amigo que

conociste hace un mes durará aproximadamente un mes más. Eso es cierto para determinadas cosas, pero no para otras.

Por ejemplo, diría que si conoces a una persona que tiene noventa años, esa persona vivirá hasta los ciento ochenta, lo cual no tiene sentido. Basándonos en nuestra experiencia, sabemos que hoy en día no es posible vivir ciento ochenta años, así que adaptamos el algoritmo de nuestro cerebro para que descarte esas predicciones. La longevidad humana se encuadra en la categoría de las cosas naturales, lo que significa que las cosas pueden crecer y vivir sólo hasta ciertos límites. Eso lo usa nuestro cerebro para hacer mejores estrategias de predicción.

Si podemos hacer buenas predicciones a partir de observaciones mínimas es porque el contexto que tiene nuestro cerebro y las experiencias previas lo hacen posible. Absorbemos esta información del mundo y lo llamamos intuición. En los casos en que no tenemos información o experiencia previas, nuestras predicciones no son buenas. Lo que proyectamos sobre el futuro revela mucho del mundo en el que vivimos y de nuestro propio pasado. Por ejemplo, para que nos vaya bien en la vida es importante desarrollar autocontrol, tener fuerza de voluntad, evitar demorar las cosas y tomar decisiones con sentido común. Pero es igualmente importante crecer en un entorno con unas referencias y experiencias buenas y sanas, ya que éstas determinarán la base que proyectamos hacia el futuro.

El cerebro también necesita decidir cuándo prestar atención a algo y priorizar para hacer buenas predicciones, usar sus propios algoritmos. La multitarea, o hacer demasiadas cosas a la vez, hace que corramos el riesgo de perder datos o de que empleemos demasiado tiempo yéndonos de un elemento a otro del contexto. Al final, si pierdes demasiado tiempo repasando el contexto y haciendo multitarea, acabarás por no obtener nada. Es como intentar cocinar demasiados platos a la vez y pasar demasiado tiempo preparando los ingredientes y los accesorios, pero sin ponerse a cocinar nada realmente, porque la preparación que requiere cada plato es diferente, y acabas deshaciendo el trabajo que ya has hecho constantemente, con lo que al final te mueres de hambre. Todos intentamos hacer las cosas lo más rápidamente posible, pero intentar hacerlo en demasiadas tareas y al mismo tiempo puede suponer un problema. Igual que un malabarista, cuando va añadiendo bolas hasta llegar a un punto en que, si añade una pelota más, sobrepasa el umbral crítico y no sólo perderá velocidad, sino que se le caerán todas las bolas al suelo.

A veces también necesitas usar un algoritmo para pensar menos para poder hacer predicciones mejores. En la era del aprendizaje automático y de la investigación sobre inteligencia artificial, la cuestión de hasta qué punto hay que dedicarse a pensar es la base de un problema llamado sobreajuste. De hecho, pensar menos puede ser una solución más sabia y un buen algoritmo. La cuestión de hasta qué punto pensar y cuántas cosas tomar en consideración es importante y juega un papel relevante en la posición que tomamos ante las dietas, el gimnasio o nuestras finanzas personales.

Un modelo demasiado complicado o al que le demos demasiadas vueltas produce soluciones que no son fiables, y que sólo son aplicables a un contexto determinado, pero no extrapolables a otro.

Por ejemplo, podemos someternos a una dieta extrema para reducir el riesgo de sufrir cardiopatías comiendo una combinación de alimentos que reduzcan el nivel de grasa corporal y aumenten el volumen muscular. Esto en principio suena bien, pero puede provocar otros problemas de salud que deriven precisamente en el problema que queremos evitar: las cardiopatías. Si decides optimizar al máximo sólo unos parámetros o consideras demasiados datos, tu capacidad para predecir el futuro puede verse mermada. Si no puedes explicarlo con palabras sencillas, no podrás comprenderlo suficientemente bien ni podrás predecir el resultado. Frenar el cerebro puede ser útil. Quizá la solución «racional» sea seguir nuestro instinto.

El cerebro está aprendiendo y adaptando sus modelos y algoritmos del mundo constantemente. El aprendizaje automático y los algoritmos de la IA también lo hacen. Una conocida categoría de algoritmos de aprendizaje automático se basa en lo que llamamos redes neuronales, que recuerdan las redes de neuronas del cerebro. Cuando el cerebro ve algo, la imagen de un amigo, por ejemplo, la señal procedente de la retina produce un resultado basado en el modelo que tiene el cerebro de tu amigo (por ejemplo, su altura, su aspecto, su forma de moverse, etc.). Si el resultado que da tu modelo cerebral no es capaz de predecir que tu amigo es efectivamente tu amigo, el cerebro calibra la información almacenada en las neuronas hasta que el error cometido por ese modelo en la identificación de tu amigo al verlo en diferentes situaciones se vuelve cada vez menor. A medida que el cerebro ve más imágenes y obtiene más datos, el modelo cerebral se va perfeccionando y comete cada vez menos errores.

Las redes neuronales se han usado en numerosas aplicaciones, desde hacer predicciones en el mercado de valores hasta conseguir que los coches conduzcan solos. De hecho, las redes neuronales como las que usa el cerebro fueron lo que usó el equipo de investigadores que ganó la competición DARPA en 2005, consistente en crear un coche capaz de atravesar el desierto conduciéndose solo, detectando la carretera con imágenes de vídeo y girando el volante en consecuencia. El equipo consiguió que el coche cruzara Estados Unidos de costa a costa con una capacidad visual muy reducida, un cerebro más pequeño que el de un gusano y con la mínima asistencia de un copiloto humano.

Hoy en día hemos llevado esas redes neuronales al siguiente nivel añadiendo varias capas de redes, como sucede en el cerebro, para que analicen una situación determinada a alta velocidad desde muchas perspectivas diferentes y hagan así predicciones muy ajustadas. Es la revolución del *deep learning*, que recurre a redes más profundas, con capas, para inventar nuevas interfaces de reconocimiento del habla en diferentes idiomas, reconocer objetos al instante, descubrir fármacos y otros logros fascinantes en diferentes sectores. Una de las ventajas del *deep learning* es su

capacidad de aprender con un número menor de datos, y de aprender por sí mismo, permitiendo por ejemplo clasificar de manera rápida las imágenes capturadas por un coche sin conductor (por ejemplo diferenciando entre bicicletas y motos con gran rapidez).

Estas redes neuronales están permitiendo avances en inteligencia artificial nunca vistos antes. Por ejemplo, AlphaZero, un programa informático desarrollado por Google DeepMind, que utiliza el enfoque generalizado de redes neuronales. El 5 de diciembre de 2017, el equipo de DeepMind lanzó resultados presentando AlphaZero, que logró en 24 horas un nivel de juego sobrehumano en Go (un juego de tablero estratégico para dos jugadores que se originó en la antigua China hace más de 2.500 años), al derrotar a los campeones del mundo, Stockfish, Elmo y derrotando a la versión anterior del mismo sistema de inteligencia artificial (que ya había ganado a los campeones del mundo 4-0), 100 a 0! AlphaZero dominó a Stockfish después de solo 4 horas de autoaprendizaje, sin acceso a libros de apertura o base de datos humanas externas. El sistema fue capaz de adquirir toda la información que ha costado a la humanidad desarrollar miles de años (la máxima capacidad en el juego Go) en unas pocas horas, compitiendo contra sí mismo, simulando millones de partidas, y desarrollando cada vez redes neuronales más complejas.

Este nuevo campo de *deep learning* que lleva los conceptos del aprendizaje del cerebro a un nivel superior está llenando titulares de periódico y se espera que mejore aún más, a medida que los ordenadores consigan procesar cada vez más datos y cada vez más rápido. De hecho, Elon Musk, consejero delegado de SpaceX y de Tesla, participa en una iniciativa que proyecta crear una interfaz cerebro-ordenador llamada Neuralink. La empresa se propone crear mecanismos que se puedan implantar en el cerebro humano, con el objetivo de permitir que los seres humanos integren software y puedan seguir el ritmo a los avances logrados en el campo de la inteligencia artificial. Estos mecanismos potenciadores podrían mejorar la memoria o permitir una relación más directa con los mecanismos de computación. ¡La inter-polinización entre ordenadores y cerebro, y su retroalimentación, no ha hecho más que empezar! La próxima generación de humanos puede ser mucho más distinta que la actual, y su diferencia con nuestra versión anterior, los chimpancés, miles de veces más grande. Las capacidades de producir nuevas mentes, cuerpos y formas de inteligencia que combinan la inteligencia artificial y el humano, pueden convertirse en la industria del futuro. Y todo esto estará determinado por grandes cantidades de datos.

La IA te cambia la vida

El conocimiento se ha ido desarrollando a través de la evolución, la experiencia y la cultura. No hace tanto que hay una nueva fuerza presente en nuestras vidas: los ordenadores. Y es algo extraordinario. Ahora hay conocimiento procedente de los ordenadores. Y eso supone un cambio enorme. El descubrimiento del conocimiento a través de los ordenadores se ha acelerado exponencialmente. En el futuro la mayor parte del conocimiento del mundo se adquirirá a través de los ordenadores. Y eso es un gran cambio.

Los ordenadores adquieren nuevos conocimientos a través de la IA y del aprendizaje automático. El profesor Domingos, en vista de lo que viene, presenta una imagen de paraíso de ciencia ficción. Plantea diversos modos en los que la IA cambiará el mundo y nuestra propia existencia.

Hoy en día, los algoritmos recomiendan libros y películas, y tu *smartphone* empieza a entender lo que dices. Todas estas cosas se consiguen con el aprendizaje automático: los ordenadores crean modelos de ti a partir del rastro de datos que vas dejando. En los próximos años, todos estos modelos se convertirán en uno y nacerá tu *álter ego digital*: un modelo en 360 grados de ti y del alma de tus registros que recorre internet constantemente en busca de las cosas que quieres —un trabajo, una cita, unas vacaciones, una casa— y que, de lo que encuentra, te hace llegar lo mejor.

El alma de tus datos interactuará por ti con los modelos de otras personas, copiándose tantas veces como sea necesario. Responderá a tus preguntas y te asesorará en todas tus decisiones. Te preguntará cómo has podido vivir sin ella, y la economía del mundo girará a su alrededor. Y eso ya está ocurriendo, con *bots* que te hablan y que personalizan sus interacciones para parecer más próximos y simpáticos.

Llegados a este punto, será posible alcanzar la «inmortalidad digital», y tendremos la capacidad de cargar nuestros cerebros con información de los ordenadores. Al transferir nuestras mentes a las máquinas, podríamos vivir para siempre, liberando el alma de nuestros datos de la debilidad de nuestros cuerpos físicos. Los conceptos de muerte y de duelo tal como los conocemos dejarían de existir.

La empresa emergente Eterni.me, con sede en Silicon Valley, tiene como objetivo conseguir que la gente pueda preservar sus «pensamientos, anécdotas y recuerdos más importantes» en un sistema inteligente artificial que podría llegar a comunicar conversando con otros una vez su creador haya desaparecido. Este avatar podría aliviar a los que han perdido a un ser querido y permitir que las generaciones futuras sepan mucho más sobre sus antepasados.

Una vez los ordenadores y los robots hagan cualquier trabajo mejor que los humanos, ¿nos quedaremos todos en el paro? A medio plazo aparecerán nuevos trabajos, de la misma forma que aparecieron durante el cambio de la era agrícola a la industrial, donde el valor de las máquinas reemplazó el valor del campo. A largo plazo no hay que preocuparse: probablemente estaremos más cerca de disponer de unos ingresos garantizados a perpetuidad, una paga mínima. Elon Musk, consejero

delegado de Tesla, también lo cree así. Él sostiene que seremos mucho más ricos que trabajando, porque la economía mundial crecerá mucho más rápidamente una vez liberada de las limitaciones que imponen nuestras capacidades humanas. La justificación económica subyacente es que, con el aumento de la automatización de las industrias —desde los transportes hasta la producción de alimentos—, habrá menos demanda de mano de obra en general, mientras que los sistemas automatizados crearán un plusvalor constante.

La gente encontrará un nuevo significado en las relaciones humanas, en ponerse al día, y los trabajos se trasladarán a sectores en los que es necesaria la presencia y la atención humanas, como la gestión, el liderazgo, la creatividad, las técnicas de asistencia o de negociación. El trabajo no servirá ya como medio para redistribuir la riqueza, y por tanto tendremos que pasar a otro modelo retributivo. Pero a corto plazo, además necesitaremos reaccionar y dar una nueva formación a los que realizan trabajos que quedarán obsoletos con la automatización para poder gestionar bien esta transición. Por ejemplo, Singapur tiene un fondo de recapitación para las personas que han perdido el trabajo por la automatización. También será necesaria la capacitación proactiva, por ejemplo, cambiando la forma en que se enseña hoy en día a los contables o a los radiólogos, porque sus trabajos cada vez los hacen menos ellos, y habrá que fomentar el acceso universal a cursos de formación en línea (MOOC) y otras herramientas educativas.

Por ejemplo, Sebastian Thrun no le ha declarado la guerra a las universidades convencionales y afirma que su objetivo no es competir con ellas. Sin embargo, en 2012 dejó su puesto de profesor de Inteligencia Artificial en la Universidad de Stanford para lanzar Udacity, una plataforma de cursos online (MOOC) que ya suma más de ocho millones de estudiantes en todo el mundo y que permite acceder a puestos de trabajo en empresas como Google con una formación de tan sólo seis meses. La gran revolución: a los alumnos que no encuentran trabajo en los seis meses posteriores a la graduación se les devuelve el importe de la matrícula, que ronda los 170 euros al mes. El foco de los cursos con más éxito es en macrodatos, inteligencia artificial, *Deep learning*, o cómo construir coches autónomos —las herramientas del ingeniero del futuro.

Con el tiempo, quizá incluso haga falta una regulación más extrema para que los humanos sean comercialmente competitivos frente a los robots o un nuevo contrato social.

Todas las tareas que puedan automatizarse se automatizarán. Tu robot te llevará en coche y te mantendrá la casa en perfecto estado. Cuando vea algo roto, encargará los recambios y lo reparará en cuanto éstos lleguen, una hora más tarde. Cuando llegues a casa al final del día, no sabrás siquiera qué es lo que ha fallado.

Recientemente, un convoy de camiones sin conductor atravesó Europa y llegó al puerto de Rotterdam. Ninguna tecnología automatizará más puestos de trabajo —ni impulsará una mayor eficiencia económica— que la de los camiones sin conductor. Hoy en día enviar un camión cargado de Los Ángeles a Nueva York cuesta unos 4.500 dólares, y la mano de obra representa el 75 por ciento de ese coste. Pero esos ahorros en mano de obra no son los únicos beneficios que aportará el uso de camiones sin conductor.

Frente a las restricciones por ley de las horas de conducción, que dictan que no se puede conducir más de once horas diarias y que hay que hacer un descanso de ocho horas, un camión sin conductor puede conducir prácticamente 24 horas al día. Eso significa que esta tecnología duplicaría el rendimiento de la red de transporte de Estados Unidos con un 25 por ciento de su coste actual.

Y el ahorro se vuelve aún más significativo si se tiene en cuenta el uso más eficiente de combustible. La velocidad óptima de crucero desde el punto de vista del máximo rendimiento del combustible es de unos 72 kilómetros por hora, pero los conductores que cobran por kilómetro recorrido van mucho más rápido. El ahorro en combustible será aún mayor cuando las flotas de camiones sin conductor adopten tecnologías de *platooning*, como la de Peloton Technology, con la que los camiones circulan en filas, formando «trenes de carretera».

El transporte en camiones representa una fracción considerable del coste de todos los artículos que compramos, de modo que los consumidores de todo el mundo notarán este cambio, traducido en precios más bajos y un nivel de vida mayor.

La revolución de la IA será diferente de la revolución industrial. La revolución industrial duró cien años, lo que representa cuatro generaciones. Fue tiempo suficiente como para que las generaciones sucesivas fueran cambiando su formación para adaptarse a los trabajos del futuro. La perturbadora tecnología que traerá la IA probablemente se implante muy rápidamente (quizá en una sola generación). Tendremos que adaptar nuestra formación en el transcurso de nuestra vida, o si no nos quedaremos sin trabajo.

Vale la pena señalar que, en ciertas situaciones, la IA aumentará el empleo. Por ejemplo, la IA está mejorando drásticamente la eficiencia en el emparejamiento de donantes de riñones con los pacientes necesitados de un riñón, lo que aumenta el volumen de trabajo de los cirujanos. Pero, por lo general, la tecnología de la IA va a dejar obsoletos muchos de los empleos de hoy en día. Aunque a mucha gente le preocupa que los robots se queden con los puestos de trabajo humanos una vez se haya desarrollado por completo la inteligencia artificial, está claro que nadie ha puesto el trabajo de «escritor» en el primer puesto de la lista de trabajos para robots. Sin embargo, ahora que un programa de IA japonés ha participado en la creación de una novela corta que ha pasado el primer examen para un premio literario nacional,

da la impresión de que no hay ningún empleo que tenga la continuidad garantizada. La novela escrita por un robot no ganó el premio final en el concurso pero ¿quién dice que no mejorará en su próximo intento?

En el campo militar, los pilotos de caza siempre han sido descritos como los mejores de los mejores. Una IA para pilotar desarrollada por un estudiante de doctorado de la Universidad de Cincinnati ha demostrado que puede superar no sólo a otras IA, sino también a un piloto de caza profesional con décadas de experiencia. En una serie de simulaciones de combate, la IA consiguió esquivar al coronel de la aviación estadounidense Gene «Geno» Lee y abatirlo en todas las ocasiones. Lee declaró que era «la IA más agresiva, ágil, dinámica y creíble que he visto nunca».

Y «Geno» no es un principiante. Ha sido estratega de guerra de las Fuerzas Aéreas, es especialista en táctica e instructor. Ha controlado o participado en miles de campañas de interceptación aire-aire como jefe de misión o piloto. En pocas palabras, este tipo sabe lo que se hace. Además, lleva luchando contra oponentes de IA en simuladores de vuelo desde hace décadas.

Teniendo en cuenta sólo las variables más relevantes, el simulador de IA puede tomar decisiones complejas a una velocidad enorme. Así, puede calcular las mejores maniobras en un entorno complejo y dinámico en un tiempo 250 veces menor que el que tarda su oponente humano en parpadear.

Tras varias misiones de combate de una hora contra el simulador de IA, Lee declara: «Me vuelvo a casa exhausto. Estoy cansado, exprimido y agotado mentalmente. Será inteligencia artificial, pero representa un desafío real para las operaciones militares actuales».

¿Te gustaría hacer un dúo al piano, pero te falta compañero? No hay que preocuparse. Lo puedes hacer igualmente, gracias al nuevo experimento interactivo de Google, el AI Duet. Básicamente tú tocas unas cuantas notas, y el ordenador toca otras notas en respuesta a tu melodía. Lo mejor de AI Duet es que toca contigo usando el aprendizaje automático y la IA, no como lo haría una máquina programada para tocar música con las notas y las normas que le han introducido en el código.

Según Yotam Mann, miembro del equipo Creative Lab de Google, AI Duet se ha visto expuesto a numerosos ejemplos de melodías. Con el tiempo, ha aprendido las relaciones entre las notas y los tiempos y genera sus propios mapas musicales basándose en lo que ha «escuchado». Esos mapas se archivan en las redes neuronales de su inteligencia artificial. Cuando le tocas una melodía al ordenador, éste compara lo que estás tocando con lo que ha aprendido y responde con lo que mejor se le ajusta a tiempo real. Eso produce respuestas «naturales», y el ordenador puede incluso producir algo para lo que no se le programó.

Y quizá acabemos yendo menos al médico. El aprendizaje automático, combinado con los datos de los secuenciadores genéticos, por ejemplo, nos permitirán crear modelos muy detallados y precisos de cómo evoluciona nuestra salud. Luego los sistemas de IA usarán esos modelos para diseñar fármacos personalizados para cada enfermedad millones de veces más rápido que en un laboratorio. Cuando se descubra una nueva infección, los sistemas de IA diseñarán un remedio ese mismo día, y tu sistema inmunitario descargará la receta de internet y empezará a fabricarlo.

Recientemente los científicos han descubierto cómo editar el ADN. Unos algoritmos capaces de aprender están creando, paso a paso, un modelo detallado de cómo emergemos de nuestros genes. Cuando completen el trabajo, pasaremos a una nueva era de medicina de precisión.

La gente no se sentirá completamente satisfecha hasta que disponga de asistencia sanitaria de alta calidad y a precio asequible en su propio barrio. Con la disminución del número de médicos que se gradúan cada año, eso no significará necesariamente ubicar a médicos en los barrios de todo el mundo. Es más probable que la tecnología llegue a todo el mundo, lo que hará posible que cualquiera se convierta el director general de su empresa de servicios de atención sanitaria. Todo el mundo podrá disponer de dispositivos en miniatura que permitirán extraer una gota de sangre, una gota de orina o una muestra de esputo y analizarlas inmediatamente para hacer pruebas de una amplia gama de problemas de salud.

Lo que antes requería una larga visita al médico, un vial de sangre y días o semanas de espera se podrá hacer en casa, con sólo una gota de sangre y una espera de quince minutos, todo ello por menos de un dólar. El diagnóstico de coste cero será una revolución de los datos y la IA en el mundo de la medicina, pero también lo serán los cirujanos robotizados que ahora mismo están en fase de desarrollo. Estas máquinas especializadas ejecutarán operaciones sencillas y repetitivas (como las de cataratas) con una precisión infalible. Si surge alguna dificultad, siempre puede recurrirse a un cirujano que puede participar desde una ubicación remota por medio de la telemedicina. Los robots cirujanos ya se usan hoy en día en todo el mundo para ayudar a los traumatólogos con las sustituciones de rodilla y otros procedimientos delicados.

Es más, la investigación de las células madre está avanzando a una velocidad impresionante. Este campo puede llegar a revolucionar completamente la asistencia sanitaria en el futuro si, tal como se prevé, se hace posible el tratamiento y la curación de enfermedades con sólo inyectar células madre en los pacientes de manera personalizada a través de IA. La asistencia sanitaria está iniciando un período de transformaciones explosivas que aumentarán el nivel de referencia de lo que se considera «buena salud» en todos los países del mundo.

¿Te da miedo la IA?

Habría que pedir a todos los estudiantes de instituto que escribieran una redacción sobre el futuro. Hay unas cuantas tendencias que muy probablemente hagan que las décadas siguientes a su graduación sean muy diferentes a la actual. En 2020 el 50 por ciento de los trabajadores en activo serán *millennials*, en 2025 serán el 75 por ciento y serán ellos los que experimenten el auge de la computación inteligente y la difusión de la infraestructura inteligente por todas partes.

Mientras Google y otras empresas trabajan para crear coches sin conductor, el sector muestra cierta preocupación. La cuestión es ésta: ¿Se sentirá cómoda la gente metiéndose dentro de uno de estos coches? ¿Tú meterías a tu hijo en un coche sin volante?

La mera idea de que los ordenadores se hagan cargo de ciertas facetas de nuestras vidas pone nerviosa a mucha gente. Pero cada mecanismo automático nuevo que entra en nuestras vidas, empezando por las puertas y las escaleras mecánicas automáticas, ha tenido que enfrentarse a este incómodo momento en que la gente se muestra escéptica e incluso hasta asustada.

Cuando se inventaron los ascensores, eran un artilugio aterrador. Pensemos en ello: estás colgado de un cable a decenas de metros de altura. Y en aquellos días había ascensoristas que tenían que operar el mecanismo para que la cabina se parara a ras de cada piso. Abrían y cerraban las puertas manualmente. Y a veces cometían errores.

Imaginemos que eres un hombre de negocios y que dices: «Vaya, pierdo el ascensor». Y corres, y das un salto para meterte en la cabina, sorteando el hueco que la separa del piso. Y consigues entrar antes de que el ascensorista cierre la puerta. Si eres un tipo atlético, todo va bien. Si tropiezas y te caes, y te quedas mitad dentro y mitad fuera de la cabina, el ascensorista no va a tener tiempo de parar. Y eso puede acabar muy mal.

Ésa no era la imagen que quería dar la industria del ascensor para su nuevo medio de transporte. Así que incorporaron puertas con topes de seguridad y freno automático. Y al final crearon el ascensor sin ascensorista. Y eso fue en 1900. Fue asombroso. Era el coche de Google de su época. Y a la gente no le gustó nada. Entraban, miraban y volvían a salir. Se daban la vuelta y buscaban a alguien para quejarse: «¿Dónde está el ascensorista?»

Pero entonces, en 1945, los ascensoristas de Nueva York se declararon en huelga. La ciudad de Nueva York quedó inmovilizada. La huelga le costó a la ciudad cien millones de dólares en impuestos no recaudados. Impidió que millón y medio de oficinistas llegaran al trabajo. Los dueños de los edificios exigieron un cambio. Y los fabricantes de ascensores decidieron que tenían que convencer a la gente para que se replanteara lo que era un ascensor.

Los primeros anuncios mostraban a niños apretando los botones y a la abuela subiendo en el ascensor para demostrar lo seguro que era. Y para los más nerviosos había una voz tranquilizadora que salía por los altavoces cuando entrabas en la

cabina: «Esto es un ascensor automático. Por favor, apriete el botón de la planta a la que desea ir». Y la misma voz te indicaba el mecanismo más tranquilizador que se ha inventado nunca: un gran botón rojo que decía «STOP». Decía: «Ha accionado la parada de emergencia. Si no es una emergencia, vuelva a pulsar el botón. Si lo es, por favor, use el teléfono».

Y estas innovaciones funcionaron. Hoy en día casi se pueden contar los ascensoristas de la ciudad de Nueva York con los dedos de las manos y de los pies. Y las lecciones que se desprenden de esto están siendo objeto de examen por los diseñadores del aprendizaje automático que ahora mismo construyen sistemas inteligentes que en un primer momento parecerán peligrosos y poco prácticos, pero que muy pronto harán nuestras vidas cada vez más fáciles.

En 1996, el Deep Blue de IBM se convirtió en el primer supercomputador que derrotó a un maestro del ajedrez, Garri Kaspárov, en una partida. Un año más tarde Deep Blue superó a Kaspárov en un encuentro completo por $3\frac{1}{2}$ – $2\frac{1}{2}$. Deep Blue derrotó a Kaspárov usando la fuerza bruta, examinando 200 millones de movimientos por segundo. Y lo más inquietante es que en las últimas dos décadas ese poder de computación no ha hecho más que volverse más brutal. En los campeonatos de ajedrez celebrados en Bilbao (España) en 2004 y 2005, un equipo de tres ordenadores derrotaron a sus oponentes humanos por $8\frac{1}{2}$ – $3\frac{1}{2}$ y 8–4, respectivamente. Pero eso fue hace dos décadas. Los *smartphones* modernos hacen que hasta Deep Blue parezca desesperantemente lento: un Samsung Galaxy S5, por ejemplo, puede ejecutar 140.000 millones de operaciones con coma flotante por segundo, una velocidad más de diez veces superior a la del antiguo ordenador de IBM.

La Ley de Moore predice que el poder computacional se duplica aproximadamente cada dos años, así que para la Super Bowl número 100, en 2066, los ordenadores podrían ser varios millones de veces más rápidos que en la actualidad. Imagínate un robot examinando un catálogo digital con billones de movimientos durante el tiempo muerto de 40 segundos entre jugada y jugada.

Jim Rushton, jefe de la división de Deporte y Ocio de IBM, predice que en los próximos años el aprendizaje automático podría ayudar a los equipos a predecir y reducir las lesiones, y a escoger a los mejores jugadores del draft. La aplicación de la inteligencia artificial a la selección de jugadas es algo inminente. Pero aunque el aprendizaje automático podría perfilar los informes de los cazatalentos que proponen jugadores, y aunque la teoría del juego pudiera ofrecer opciones circunstanciales, todo esto no son más que herramientas que puede emplear un entrenador. Y en 2066, ese entrenador seguirá siendo humano. Probablemente.

Muy pronto el aprendizaje automático dará lugar a una inteligencia sobrehumana. A partir de las noticias que nos llegan cada día de los progresos en IA es fácil tener la impresión de que los ordenadores están a punto de ver, hablar y razonar como

nosotros, tras lo cual muy pronto nos dejarán atrás. Desde luego hemos llegado muy lejos en los primeros cincuenta años de inteligencia artificial, y el aprendizaje automático es la razón principal de sus recientes éxitos, pero aún nos queda mucho camino por recorrer. Los ordenadores pueden ejecutar muchas tareas limitadas muy bien, pero aún no tienen sentido común o consciencia, y nadie sabe exactamente cómo enseñarles a adquirirlo.

En los próximos años veremos cómo ayuda la inteligencia artificial al ser humano y viceversa, creando una mezcla de inteligencia natural y artificial. Veremos nuevas técnicas de análisis de datos basadas en la inferencia de necesidades e información, y no en la búsqueda del «conocimiento perfecto». Los nuevos sistemas de inteligencia aumentada usarán una cantidad mínima de información digital, derivada de la enorme cantidad de datos recogidos, y de ella inferirán lo que se necesita en cada situación y contexto, para asistir —no reemplazar— a la inteligencia humana.

Con demasiada frecuencia la respuesta a esta cuestión recuerda la trama de un *thriller* de ciencia ficción. A la gente le preocupan que los avances en IA incluyan la «singularidad», ese punto de la historia en que la IA sobrepasa a la inteligencia humana, llevando a una inimaginable revolución e intromisión en los asuntos de los humanos. O quizá se pregunten si podremos controlar la inteligencia artificial o si ella nos controlará a nosotros, convirtiéndonos en cibernéticos.

Son cuestiones interesantes, pero no son urgentes. Tienen que ver con situaciones que quizá no se planteen en cientos de años, si es que llegan a plantearse. De momento nuestras mejores herramientas de IA (como el programa AlphaGo de inteligencia artificial de Google que recientemente venció al mejor jugador de Go del mundo) no tienen forma de adoptar una dimensión «general» y hacer razonamientos con sentido común, adquirir conocimientos en dominios diversos, sentir, expresar, comprender las emociones, etc.

Eso no significa que no haya nada de lo que preocuparse. Al contrario, los productos de IA que existen actualmente van mejorando más rápidamente de lo que cree mucha gente y prometen transformar radicalmente nuestro mundo, no siempre a mejor. Son sólo herramientas, no una forma de inteligencia con la que haya que competir. Pero darán un nuevo significado al trabajo y al concepto de riqueza, provocando unas desigualdades económicas sin precedentes e incluso alterando el equilibrio de poder global.

Los «empleos» no quedarán eliminados por completo con la IA y la automatización, pero las «tareas» sí. La mayoría de empleos de la gente se componen de una serie de tareas, de modo que la automatización puede ir afectándoles con el tiempo, no de golpe. Es importante darse cuenta de que esto no es una tendencia nueva, sino más bien una extensión lógica de lo que ha ido sucediendo durante siglos. Los sociólogos siempre han descrito el progreso tecnológico como la elaboración de

herramientas que al desarrollarse permiten cubrir las necesidades fundamentales cada vez con mayor eficiencia y con menos esfuerzo humano. ¿Cuál es la característica de la ciudad moderna? Que ya no tenemos que ir al río a buscar el agua, ni plantar y cosechar los campos.

Con la división del trabajo y la industrialización, la sociedad se vuelve cada vez más eficiente y saca más partido al trabajo. La automatización industrial ha ido avanzando, y con ello los trabajos más duros los han ido asumiendo las máquinas. El aprendizaje automático y la IA no son más que eso mismo. ¿Qué tiene de particular, pues, la situación actual? Yo diría que, quizá por primera vez, la creciente automatización del trabajo intelectual afecta al concepto básico que distingue a distintos tipos de trabajo. Incluso las personas con niveles de formación más avanzados (por ejemplo, abogados, médicos, etc) aún no se han hecho a la idea de que quizá ellos también van a verse desplazados con ese proceso, y que quizá pierdan categoría o privilegios. Históricamente, las grandes disrupciones han llegado de forma discontinuada —por salto y no por evolución—, cuando una masa crítica de personas se dan cuenta de que las ventajas económicas de una nueva tecnología a nivel mundial acabarán imponiéndose en el mercado. Y ahora, con la IA, se nos da mejor que nunca ver y distinguir esas oportunidades.

La IA y los macrodatos

Simplemente recopilando o teniendo acceso a conjuntos de datos mayores no obtendremos resultados. No podemos mantenernos al día con el tsunami de datos de más de 100 terabytes al día, entre vídeo en streaming, texto, imágenes y voz. Por otra parte, capturar la cantidad de éstos almacenados en papel y otros medios resulta imposible. No estamos preparados para la extracción de conocimientos y para la toma de decisiones cada vez más rápida que nos exigirán nuestros clientes y nuestros mercados, nuestro planeta, con el fin de conseguir un alto rendimiento, una ventaja competitiva, y probablemente para la simple supervivencia. La convergencia de estos dos frentes acelerará la innovación: la consecuencia puede ser una mayor rapidez en el descubrimiento de los fármacos, coches sin conductor más seguros, fuentes de energía alternativas y muchas otras innovaciones aún por llegar.

El uso del aprendizaje automático, los sistemas expertos de IA y la analítica combinada con los macrodatos es la evolución natural de lo que hasta ahora han sido dos disciplinas diferentes. Ahora están convergiendo.

Capturar macrodatos resulta útil, por ejemplo, para identificar tendencias o patrones en el comportamiento de clientes o empleados. Pero lo importante es hacer posible llegar a extraer conclusiones, incluso de forma automática, para encontrar los medios óptimos para ser productivos, útiles o incluso para resolver problemas — desde enfermedades hasta nuevos modelos de negocio o incluso la lucha contra el

hambre—. La IA se usará para extraer conclusiones, decidir mejores resultados y posibilitar una toma de decisiones más rápida a partir de fuentes de enormes macrodatos, y eso dará forma al futuro de los negocios, de la prosperidad y del desarrollo planetario. En un mundo en que hay macrodatos por todas partes, será la IA la que impulse la extracción de conclusiones, la monetización de los datos con un objetivo.

¿Qué deparará el futuro a la «inteligencia de los macrodatos», producto de la convergencia de los macrodatos y la IA? ¿Qué desafíos aparentemente insuperables podremos afrontar, pese a que aún no consigamos comprenderlos bien? ¿Mejores empleos, un entorno más sostenible, una economía más sólida, un mundo más seguro, el fin del cáncer? Tenemos que pensar así si queremos dar forma al futuro.

Cada vez se imponen más el uso de sistemas de apoyo a la toma de decisiones y la toma de decisiones de forma autónoma. Aún hay muchas cuestiones que afrontar; en este campo estamos en una fase temprana. Los sistemas de aprendizaje automático que determinan qué macrodatos identificar y usar de momento requerirán la participación del ser humano para dirigirlos, pero no así en un futuro cercano. Vemos cómo va arraigando la IA en el campo de la asistencia sanitaria, en el derecho, en las finanzas, en la gestión de inversiones, en el comercio de valores, en la publicidad, en la seguridad, en la defensa y en los servicios financieros, donde sólo los macrodatos no bastan. Necesitamos la *Big Data Artificial Intelligence* (inteligencia artificial de datos masivos).

Cuando le digo a la gente que dirijo un equipo de investigación de *big data* e inteligencia artificial, es inevitable que la conversación desemboque en la pregunta: «¿Por qué construyes máquinas que nos quitarán el trabajo a todos?». Sin embargo, lo que a menudo se pasa por alto es la ética de la IA.

Es esencial que los usuarios confíen en los sistemas de IA, o se corre el peligro de que la sociedad no los acepte bien. Por ejemplo, si un coche sin conductor no se gana la confianza del usuario, es poco probable que nadie quiera usarlo. Para crear algoritmos de IA dignos de confianza tenemos que hacer que sean seguros y precisos. También tenemos que hacer que la IA sea justa. Tenemos la responsabilidad de hacer que sea justa. Eso significa eliminar cualquier sesgo indeseable. Desgraciadamente, en la actualidad tenemos muchos ejemplos de sistemas de IA sesgados o injustos. Por ejemplo: a) los sistemas usados para calcular el riesgo de reincidencia en el sistema penal de Estados Unidos discriminan a los afroamericanos; b) el sistema de reconocimiento de voz de Google ha demostrado un rendimiento sistemáticamente mejor con voces masculinas que con voces femeninas; y c) la mayoría de los coches sin conductor tienen el sesgo de que su aprendizaje parte de los datos de California. Por último, la IA requiere honestidad: los algoritmos deben ser transparentes e interpretables (no cajas negras). Deberían poder darnos explicaciones de las

decisiones que toman. Eso va a convertirse en una preocupación comercial significativa en la Unión Europea con la nueva Regulación General de Protección de Datos (GDPR) en 2018. Es necesario que los usuarios tengan el derecho a recibir una explicación lógica de cómo usa nuestros datos personales un algoritmo. Y para ello serán necesarios avances en la transparencia algorítmica, y que se pueda explicar por qué funcionan de un modo determinado y no de otro los algoritmos de la IA.

Sería un error ver todas estas tendencias —como la convergencia de los macrodatos y la IA— como hechos diferenciados. Eso eliminaría las ventajas competitivas que ofrecerá la convergencia. Tenemos que pensar de otro modo, es mejor la coevolución que cualquiera de las tendencias que representan un sector o un mercado por separado. Es importante reconocer la conectividad de las tendencias.

En lo relativo a los macrodatos y la IA, creo que estamos en una fase inicial. Pero en un futuro próximo no habrá macrodatos sin IA, y eso puede ser dentro de un minuto o dentro de un año.

No es ningún secreto que la inteligencia artificial (IA) y el aprendizaje automático han avanzado tremendamente en la última década; sin embargo, en algún punto intermedio entre la mejora de los algoritmos y el aumento de velocidad de los procesadores informáticos está la base del rendimiento máximo —y de la obtención de mejores resultados— por parte de los sistemas de ingeniería. El problema, ahora mismo, es que los expertos de IA pasan demasiado tiempo preprocesando código y datos, iterando modelos y parámetros. Cada paso del camino requiere demasiado trabajo o demasiada computación. Esto tiene que cambiar, y para ello va a haber que rediseñar chips de computación.

Actualmente ya hay equipos en empresas como Nvidia, Microsoft, Baidu, Google, Facebook, Amazon, e Intel afrontando directamente este reto. Están experimentando, desarrollando e incluso poniendo en práctica nuevos diseños de chip, interconexiones y sistemas para procesar conjuntos de macrodatos a toda máquina y llegar donde no han llegado nunca la IA, el deep learning y el aprendizaje automático. Estos avances podrían tener un gran impacto en los próximos años —incluso un efecto revolucionario— en muy diversos campos: la conducción automática, el descubrimiento de fármacos, la medicina personalizada, los asistentes inteligentes, la robótica, el análisis de macrodatos, la seguridad informática y mucho más. Podrían aportar un procesamiento mejor y más rápido para tareas relacionadas con el habla, la visión y la búsqueda contextualizada.

Los chips especializados pueden mejorar significativamente el rendimiento de la IA y los macrodatos porque incluyen todo lo necesario específicamente para la tarea del momento y nada más. Sin embargo, hay retos que superar. Serán necesarios nuevos diseños de chips, nuevos chips especializados para la computación de la IA. Eso supone mejoras en las GPU (Unidades de Procesamiento Gráfico), que se han convertido en el elemento de apoyo más habitual del *deep learning* y de la IA en los últimos años, así como un trabajo en otras tecnologías, como el de las matrices de

puertas programables (FPGA), las unidades de procesamiento de tensor (TPU) y otros sistemas y arquitecturas de chips que encajen con requisitos específicos de la IA y los macrodatos. Estas iniciativas permitirán alcanzar el nivel de rendimiento y de eficiencia energética necesarios para crear inteligencias artificiales con un nivel de precisión que hoy no es posible.

Y es que el futuro nos reserva algunos logros realmente emocionantes. La IA y los macrodatos, una vez aplicados, tienden a adaptarse específicamente a los datos y los requisitos de la tarea del momento. Pero ¿hasta qué punto se pueden reutilizar los algoritmos de IA para diferentes tareas? ¿Pueden reutilizarse los conocimientos de una tarea para otras? El aprendizaje por transferencia es un sector emergente que abstrae el know-how de la IA e intenta reapplicarlo en otras tareas. El aprendizaje por transferencia supone el auge de la IA colaborativa.

¿Eres padre o madre y quieres que tu hijo de ocho años aprenda a tocar el violín? Vale. Pero ¿qué tiene que ver eso con la inteligencia artificial? Bienvenidos al mundo del aprendizaje por transferencia. En experimentos recientes se ha observado que los bebés —a partir de nueve meses— que aprenden música pueden mejorar significativamente muchas de sus otras funciones cognitivas, como la futura adquisición del lenguaje. Los niños que aprenden a tocar un instrumento de pequeños desarrollan mayores habilidades de aprendizaje verbal y lingüístico que los que no, porque adquieren unas habilidades de representación de sonidos y modifican su conectividad cerebral.

En el caso de los adultos, un buen ejemplo es el aprendizaje de un idioma. ¿No has notado que cuantos más idiomas aprendes, más fácil resulta? Por supuesto, es mucho más fácil aprender italiano cuando ya sabes español, pero eso se debe sobre todo a las similitudes en el vocabulario. Pero hay cosas menos obvias... ¿Por qué nos cuesta menos aprender ruso cuando ya sabemos inglés y árabe, idiomas que no comparten ninguna similitud en el vocabulario? Probablemente porque: 1) hemos integrado en el cerebro principios muy generales y estructuras abstractas comunes a diversos idiomas, y 2) hemos aprendido... a aprender.

Cuanto más estudias matemáticas, más desarrollas la lógica. Aprenderás más rápido en nuevos campos en los que es importante el razonamiento lógico, como la filosofía.

¿Qué conclusión podemos sacar? Aprender una cosa en un dominio te puede ayudar a aprender más rápidamente muchas otras cosas en diferentes dominios, a usar la lógica para combinar estructuras abstractas comunes. Exacto: esta idea tenía que ser explorada por la IA, y se le llama aprendizaje de transferencia. ¡Probablemente sea la próxima revolución de la IA!

Probablemente pienses que esto sólo es aplicable a los humanos. ¡No! También es perfectamente válido para las máquinas. Los algoritmos de traducción más populares,

como los de Google Translate, son muy sensibles a este fenómeno. De hecho, si quieres entrenar un algoritmo así para que traduzca, por ejemplo, del inglés al francés, también tendrías que entrenarlo para que traduzca de español a ruso, o de italiano a alemán. ¡Y te darías cuenta de que la calidad de la traducción inglés-francés ha mejorado mucho! Esto es asombroso desde el punto de vista filosófico: significa que empiezan a observarse fenómenos cognitivos elevados en los sistemas de IA más recientes.

El aprendizaje de transferencia también puede ayudar a los sistemas de IA para que aprendan de un número de casos reducido, tal como explica Gilles Wainrib. Probablemente sea la historia de aprendizaje por transferencia más usada actualmente, y uno de los motivos ocultos por los que tiene tanto éxito el *deep learning* de la IA. Supongamos que quieres crear un algoritmo capaz de diferenciar un cáncer de pecho de uno de pulmón a partir de pruebas de imagen (como las imágenes patológicas o las resonancias). En principio, los algoritmos tienen una gran capacidad de reconocimiento de imágenes y deberían ser la mejor opción. Para usarlos serían necesarias entre 100.000 y un millón de imágenes, con el fin de obtener un alto nivel de precisión. Pero en el caso de la mayoría de aplicaciones médicas esto podría ser un problema enorme, ya que la mayoría de conjuntos de datos que pueden reunir los médicos y los investigadores no suelen tener más de un centenar de imágenes como máximo.

¿Y si usáramos el conocimiento de otros conjuntos de datos más abundantes para acelerar el proceso? ¿Y si pudiéramos importar el conocimiento de la identificación de imágenes de gatos, camiones u ordenadores y transfiriéramos ese conocimiento a otros algoritmos de IA para que así pudieran reconocer más rápidamente tumores malignos?

En la IA, el aprendizaje por transferencia será de enorme importancia en el futuro, ya que así podrá pasarse el conocimiento de una tarea a otra, del mismo modo en que nuestro cerebro usa conocimientos y habilidades aprendidos en el pasado para resolver rápidamente problemas nuevos. Esto adquiere una relevancia especial en campos en que haya pequeños conjuntos de datos —en oposición a los macrodatos—, pero aun así queremos que la IA funcione bien. Es especialmente importante en el contexto de los estudios médicos, ya que la tendencia actual es la de recopilar cada vez más datos de cada paciente, aunque el número de pacientes con anotaciones claras sigue siendo bastante pequeño en la mayoría de casos.

Con el aprendizaje por transferencia se hace posible proponer algoritmos «preentrenados» para problemas determinados con el fin de potenciar el poder de muchos sistemas de IA. Es un recurso que tiene un potencial tremendo y que contribuye a un nuevo sistema emergente de inteligencia artificial global y cooperativa, en la que los algoritmos de IA comparten sus secretos, se ayudan entre sí, aprenden entre ellos y aceleran la resolución de problemas importantes y los descubrimientos médicos.

4

La IA para el liderazgo y la creatividad

Hacer predicciones es una tontería; tú disfruta del partido.

JOHAN CRUYFF

La Revolución Industrial tuvo lugar cuando los humanos superaron las limitaciones de su poder muscular. Ahora nos encontramos en los estadios tempranos de que pase algo similar, pero en cuanto a nuestra capacidad mental, incrementándola infinitamente en virtud de las tecnologías digitales. Y esto tendrá implicaciones en nuestra capacidad de liderazgo, de crear organizaciones, equipos, y sistemas de cooperación que superen los actuales, y que creen nuevos modelos organizacionales, de liderazgo e influencia.

No hay mejor ejemplo que demuestre el impacto de diferentes estilos de liderazgo, que los deportes en equipo. Los entrenadores, necesitan ser capaces de poder coordinar propuestas colectivas, sentenciando el futuro de los equipos en la competición para bien o para mal. El objetivo es exponer por convencimiento y no por imposición, modelos de liderazgo y cooperación que hagan mas fácil conseguir el éxito contra el contrario. Y la inteligencia artificial y el poder de los datos esta permitiendo revisitar el liderazgo en el siglo XXI y transformarlo para siempre.

En un partido de futbol, el alma del partido es la estrategia y el liderazgo, y ambos se pueden interpretar y evolucionar a través de los datos, la inteligencia artificial y los algoritmos.

Yo pasé muchos años trabajando para hacer que internet fuera más rápida y construyendo sistemas distribuidos y redes colaborativas en algunas de las entidades más innovadoras y pioneras del mundo, como Microsoft, Bell Laboratories, Telefónica o la Universidad de Columbia. Tuve el privilegio de formar parte de una comunidad de investigadores, científicos e ingenieros que se dedicaban a diseñar, construir y mantener algo hasta entonces nunca visto.

Pero a través de conversaciones, encuentros casuales —y a veces, simples ideas que se me han ocurrido— me he dado cuenta de que el conocimiento que han adquirido los científicos sobre datos e IA y la experiencia que hemos ido adquiriendo durante el vertiginoso nacimiento de internet hoy en día se puede aplicar a otras disciplinas, con resultados extraordinarios. No se trata sólo de aplicar puntos de vista diferentes, sino herramientas específicas, técnicas de datos y procesos desarrollados para un campo, adaptados a otro.

Yo me crié en España, donde se publican cinco diarios dedicados exclusivamente al fútbol. Aun así, yo nunca he sido un tipo muy deportivo. Cuanto tenía doce años

un caricaturista me hizo un retrato y creo que capturó perfectamente mi actitud con respecto al fútbol, al mostrar una pelota que entraba tranquilamente en la portería mientras yo estaba apoyado en el poste, con la cabeza en otras cosas. Todo aquello cambió, no obstante, cuando llegué a Barcelona, en plena era del *tiki-taka*.

Cuando volví a Barcelona, el fútbol que se jugaba en el Camp Nou estaba en su zénit. El Barça marcaba cuatro o cinco goles por partido con una facilidad raramente vista anteriormente, y los equipos rivales no sabían siquiera por dónde empezar a defender. Pero no me interesaba. Eso, hasta que fui al campo y vi unos cuantos partidos. Ver a los jugadores en acción me despertó la curiosidad, y tuve una revelación: yo sabía algo de aquello aunque nunca creía haberlo sabido. Sabía lo que estaba ocurriendo, y podía deducir lo que ocurriría a continuación.

E incluso cuando el Barcelona y su estilo *tiki-taka* encontraron quien les plantara cara, pensé que podía imaginar el motivo. Más aún: cuándo la gente empezó a decir que no habría más revoluciones en el fútbol, su estrategia y liderazgo de equipo, yo sabía que probablemente estaban equivocados.

Liderazgo y fútbol: el colectivismo y el deporte

Yo cursé mis estudios de ingeniero en telecomunicaciones en Pamplona (la ciudad donde se desarrolla la trama central de Fiesta, la novela de Ernest Hemingway) y luego me doctoré en redes en el Instituto Federal de Tecnología de Lausana, antes de trasladarme a Estados Unidos para diseñar redes entre pares y sistemas de internet a gran escala. En Pamplona conocí al profesor Luis Mercader, científico catalán que se crió en Moscú y que dio clases de radiocomunicaciones por satélite en la Universidad de Tecnología de Moscú, pero que posteriormente volvió a España, donde fue nombrado profesor emérito.

Fue en Pamplona donde tuve mi primer encuentro con las redes y con la ciencia, y también donde tuve la oportunidad de oír al profesor Mercader que hablaba de su hermano, Ramon Mercader, y también comprendí cómo había ayudado a los soviéticos a diseñar y planificar sus redes de satélites, para después ir a dar clase de radiocomunicaciones en la universidad de ingeniería donde yo estudiaba. Ramon Mercader, el hermano de mi profesor, fue el agente soviético que asesinó al revolucionario marxista ruso Leon Trotski en 1940, en México, cuando Trotski se opuso a la burocracia estalinista de la Unión Soviética y sugirió reformas en la estructura política.

En el año 1930 Josef Stalin impuso el colectivismo en la Unión Soviética, instaurando granjas colectivas en lugar de las privadas. La premisa era que con las granjas colectivas aumentaría la producción de alimentos para la población urbana, y que proporcionarían una solución más eficiente a la crisis de la distribución agrícola.

Este esquema económico cooperativo produciría mayores cosechas y mejoraría los ingresos del Estado, frente a la agricultura privada, considerada menos eficiente.

Stalin estaba convencido de que el concepto de la «unidad completa» tenía prioridad sobre el individuo, y de hecho la producción aumentó hasta las casi 100 millones de toneladas en 1937 a pesar de la intensa resistencia de los granjeros —más de 3 millones de kulaks fueron asesinados durante este período, y muchos granjeros decidieron quemar sus cosechas y matar a sus animales en lugar de entregarlos al Estado.

Más de cincuenta años después, un hombre que creció bajo la influencia del colectivismo soviético cambiaría la imagen del fútbol para siempre. En 1986, el Atlético de Madrid jugaba la final de la Champions League contra el Dinamo de Kiev, que acababa de ser escenario de la revuelta ucraniana contra Rusia. El Atlético de Madrid llevaba un gran registro, ya que había ganado todos los partidos de la competición, y sus resultados como visitante eran aún mejores que los del Real Madrid. Luis Aragonés, entrenador del Atlético, había construido un equipo sólido, pero aún no había conseguido conquistar el trofeo europeo.

En el otro lado estaba el Dinamo de Kiev, que también venía de una temporada de éxito. El Dinamo de Kiev estaba dirigido por Valeri Lobanovski, entrenador de cuarenta y siete años que diecisiete años antes había estado a punto de dejarlo todo para convertirse en fontanero, pero que al final acabó ganando una medalla de oro en matemáticas. El resultado final fue de 3-0, una clara victoria para el Dinamo, pero la historia recordará este partido por el segundo gol que marcó el Dinamo, posiblemente uno de los goles más bonitos marcados nunca y sin duda uno de los más importantes en la historia de la táctica del fútbol.

Durante el dominio soviético, Lobanovski había sido un gran jugador de fútbol, a las órdenes de Víktor Maslov, el gran pionero del juego de presión, táctica utilizada para recuperar la posición rápidamente. Lobanovski vivía en Kiev, que era el centro de la incipiente industria informática. El primer instituto cibernético se creó allí mismo, y muy pronto se convirtió en pionero en campos como los sistemas de control, la inteligencia artificial y los modelos matemáticos. Además, la URSS estaba a la vanguardia de la investigación científica y computacional con su programa espacial. Antes de convertirse en jugador de fútbol y más tarde en entrenador, Lobanovski había estudiado ingeniería y había interactuado con diversos científicos y estadísticos. Él veía el fútbol como un sistema de 22 jugadores, en el que 11 tenían que superar a los otros 11 para ganar (lo cual imponía una serie de limitaciones a las posibilidades de movimiento de los primeros 11), y fue pionero en el uso de la tecnología informática en el entrenamiento.

Lo más interesante es que Lobanovski pensaba que la eficiencia de un sistema de fútbol dependía en gran medida en la interacción entre los jugadores, y que esta eficiencia podía aumentarse en gran medida pensando en el grupo, en lugar de

centrarse en cualquiera de los jugadores individualmente. Hasta entonces, el fútbol se había planteado como un juego de «ataque total»: patada y carrera.

Hasta entonces las habilidades individuales, como el regateo y la carrera con el balón, eran consideradas el mejor modo para abrumar al contrario, y raramente se veían pases entre compañeros. Lobanovski, en cambio, tenía la convicción de que en el fútbol no eran tan importantes los jugadores individuales como la estructura y las conexiones entre los jugadores y la red que creaban, que podía analizarse con datos y optimizarse con algoritmos. Él veía un vínculo entre algoritmos, pases y victorias. Además, al igual que Stalin, pensaba que el colectivismo era mucho mejor que el talento individual.

Eso se hizo evidente en aquella final europea entre el Atlético de Madrid y el Dinamo de Kiev. El segundo gol se convertiría en el ejemplo más claro de que la nueva teoría de algoritmos iba a cambiar la historia del fútbol para siempre, y de que la teoría de Lobanovski y su tecnología informática funcionaban. Aquel segundo gol no recurrió a los regateos individuales, que eran la norma hasta entonces: participaron tres jugadores, en un contraataque rápido en el que dos de ellos no dispararon a puerta, aunque tuvieron ocasión para hacerlo, sino que decidieron pasar a un tercer compañero que tenía más posibilidades de marcar, en una jugada de gran belleza y fluidez que ponía en evidencia la maquinaria táctica del equipo.

En paralelo, en los Países Bajos, Rinus Mitchel estaba desarrollando un concepto similar al de Lobanovski, pero de forma intuitiva y orgánica, en lugar de usar la tecnología informática. Él ideó el concepto de «fútbol total»: hasta entonces, el fútbol se basaba en jugadores especializados en una tarea determinada (por ejemplo, atacar o defender) y que ocupaban un área muy definida del campo.

En el fútbol total, cualquier jugador podía desempeñar el papel de otro jugador (salvo el portero). Un jugador que abandonaba su posición podía ser sustituido por cualquier otro jugador de su equipo.

Este algoritmo creaba un modo de juego y de liderazgo que podía parecer ilógico: los jugadores no se apoyaban unos a otros acercándose entre sí, sino alejándose. Era una bonita visión del fútbol, nada fácil de ejecutar, pero que cambió el modo de jugar al fútbol para siempre. Dado que, teóricamente, cualquier jugador podía jugar en la posición de otro, el entrenador no necesitaba tener el jugador idóneo en cada posición para desarrollar una táctica determinada, sino que le bastaba con tener «un» jugador en una posición determinada, con lo que aumentaba la velocidad y la versatilidad del equipo.

El fútbol total holandés fue un sistema de liderazgo fluido desarrollado a partir del liberalismo de finales de los años 1960 en Ámsterdam, frente a la influencia soviética en la planificación de Lobanovski: los cambios en las sociedades a menudo discurren en paralelo a cambios en diferentes aspectos de la política, el arte, la música y el deporte, ya que todos ellos son producto de una misma sociedad. Los holandeses, en lugar de separar el juego entre ataque y defensa, y de especializar a los jugadores

en una u otra función, decidieron que cualquier jugador que estuviera en el campo podía hacer ambas cosas. Ya no se trataba de ver qué sabía hacer cada jugador, sino de asegurarse de que había un jugador en un lugar preciso del campo para que hiciera lo necesario.

Esto requería dominar muchas técnicas de entrenamiento (por ejemplo, correr atrás tras el pase, en lugar de adelante), además de ser capaz de mantener la posición mirando alrededor y comprobando la posición de los otros jugadores para saber que se está bien situado. De media los jugadores tienen el balón tres minutos por partido, así que lo más importante es lo que haces en el campo durante los 87 minutos en que no tienes el balón.

Johan Cruyff fue el más famoso exponente de este sistema, un jugador que adaptaba sus movimientos y cambiaba de puesto para que posiciones tácticas fundamentales siempre estuvieran cubiertas. «Si diseñas un equipo escogiendo al mejor jugador para cada posición, acabarás teniendo no un 11 fuerte, sino 11 «unos» fuertes», solía decir. El espacio y la creación del juego eran básicos para el concepto del fútbol total. Johan Cruyff siempre hablaba de dónde había que correr y dónde había que quedarse quieto, y de cuándo debía mantener la posición el equipo, en lugar de hablar de qué hacer con el balón. El algoritmo de cambiar constantemente de posiciones, que acabaría conociéndose como «fútbol total» surgió gracias a esta conciencia espacial. «El fútbol sencillo es el más bonito. Pero jugar un fútbol sencillo es de lo más difícil», explicaba Cruyff.

Gracias al colectivismo de Stalin, las granjas de la Unión Soviética acabaron siendo muy productivas, aunque ello tuviera un coste muy alto en vidas y en sufrimiento. Impusieron el colectivismo en las granjas privadas, creando una economía más global y eficiente. No obstante, tal como descubrió más tarde Elinor Ostrom (ganadora del Nobel en 2009), los soviéticos no tenían un sistema colectivo perfecto. Con imágenes por satélite se determinó que las praderas gestionadas con el colectivismo en China y Rusia eran menos productivas que las praderas de Mongolia situadas en terrenos similares.

La diferencia era que los mongoles habían mantenido una tradición milenaria por la que el trabajo colectivo no se imponía a través de estructuras rígidas creadas por las altas esferas del gobierno, sino que el liderazgo surgía de unas normas de cooperación definidas por las diferentes tribus y que habían evolucionado a lo largo de los años, integradas en su cultura y sus tradiciones, lo que mantenía las praderas fértiles y productivas. Aparecieron ejemplos similares en otros grupos que habían creado sus propias normas y formas de interacción, sin necesidad de imponer estructuras de liderazgo o formaciones: es el caso de los pastos suizos o las redes de riego de España o de Filipinas.

La gran dimensión de la investigación de Elinor y de su análisis, basados en datos contrastados y en la teoría de los algoritmos, hacía que sus novedosos hallazgos resultaran irrefutables. Por otra parte, apuntaban a una nueva visión de la interacción entre los seres humanos, el liderazgo y los sistemas, que también podían ser de utilidad en el fútbol: proporcionaban un argumento para mirar más allá de los límites de los recursos comunes, estructurados jerárquicamente y gestionando el liderazgo de un modo tradicional. Planteaba la necesidad de nuevos algoritmos de gobierno y liderazgo basados en la confianza entre los seres humanos y en interacciones y normas definidas por los individuos, en los que un gobierno no oficial proporcionaba mejores resultados que otros gobiernos muy estructurados. Es más, generaba los incentivos necesarios para que cada persona interactuara de buen grado y rindiera más.

En 2008, Pep Guardiola, entrenador del FC Barcelona, había dirigido 242 partidos, había ganado tres títulos de liga seguidos, en dos ocasiones acompañados de la copa de Europa, y había creado un modo elegante y estético de jugar al fútbol que había robado el corazón y despertado la imaginación de los aficionados y que hacía que los jugadores estuvieran siempre deseando jugar. La belleza desplegada por Messi, Xavi, Iniesta y el resto del equipo —que veían liberada toda su creatividad— llevó a un juego de pases de una audacia y una precisión técnica increíbles.

Las normas algorítmicas y el liderazgo instaurado por Lobanovski, Rinus Mitchel y Cruyff habían sido un gran avance para que los jugadores jugaran en equipo, como una red, y para que cada uno de ellos fuera extremadamente versátil. Aun así, la estructura del equipo durante los partidos seguía basándose en formaciones fijas, como el esquema de rombo 4-4-2, en pirámides, o las estructuras de tres fases, en las que el portero pasaba la pelota a los defensas, que a su vez la pasaba a los centrocampistas, y éstos la pasaban a los delanteros.

Guardiola, por su parte, introdujo un estilo de juego con un algoritmo completamente libre (probablemente equivocado) conocido como *tiki-taka*. A pesar de que en ocasiones podía usarse para mantener la posesión sin atacar, el *tiki-taka* iba más allá.

El modo en que empezó a jugar el FC Barcelona se apartaba radicalmente de la idea tradicional de las formaciones en tres niveles, de las jerarquías o de las pirámides, pasando a un concepto de liderazgo caracterizado por unas estructuras de equipo dinámicas, prácticamente autoorganizadas, moviendo el balón por diferentes canales y siguiendo diferentes patrones, y manteniendo la posesión con pases cortos, más difíciles de interceptar que un pase largo.

Compartía con la idea holandesa del fútbol total tanto el intercambio de posiciones como el uso de la posesión para controlar el partido. Pero se alejaba del sistema holandés en que todo quedaba subordinado al pase, hasta la estructura del

equipo; a una serie de patrones algorítmicos de pases que surgían de una serie de interacciones fluidas entre los jugadores gobernadas por una interpretación del juego como un sistema basado en una estructura de equipo autogenerada y en la cooperación dinámica. Este estilo implicaba movimientos en todas direcciones e intercambios de posición: «bailar» y mover el balón en complicados patrones producto de las interacciones entre los jugadores.

Con el *tiki-taka*, la pirámide, la estructura y la jerarquía desaparecieron. La estructura del equipo surgía de los pases, en lugar de ser la estructura la que determinara cómo se pasaba el balón, como se había hecho hasta entonces. Desde aquel momento, el *tiki-taka* se ha asociado a la clase, a la creatividad y a la cooperación entre jugadores, del mismo modo que las tribus mongolas descubiertas por Elinor Ostrom habían definido sus propias formas de cooperación, tan exitosas durante milenios.

Curiosamente, en España el *tiki-taka* apareció simultáneamente al movimiento de liderazgo social de los «indignados» (y luego al «Occupy Wall Street», en Estados Unidos), que impulsó un movimiento ciudadano contra el sistema político establecido, proponiendo un modo cooperativo, más orgánico y ascendente, de organizar la sociedad, con gobiernos menos centralizados y jerárquicos y una organización más transversal. España ganó el Mundial jugando con siete jugadores del FC Barcelona, y en 2013 Guardiola dejó el club y se fue al Bayern de Múnich.

Al igual que en el caso del fútbol total, los algoritmos del *tiki-taka* no eran fáciles de poner en práctica. De hecho, muchos de los jugadores no se adaptaban bien a este tipo de juego. Los futbolistas tenían que sacrificar su juego individual para participar en el estilo de juego fluido y cooperativo del equipo. Para adoptar ese estilo de juego era necesario que todo el equipo fuera perfectamente consciente de los incentivos que suponía. Por ello fue fundamental la formación de los más jóvenes en la academia de fútbol del FC Barcelona, La Masia, y fichar a jugadores que fueran capaces de adaptarse al estilo del equipo.

Muchos años antes, en Kiev, donde Lobanovski había estudiado y había tenido la inspiración de usar la tecnología informática para entrenar, el Instituto de Cibernética de Kiev empezó a planificar una drástica reforma del sistema económico soviético basado en la computarización de finales de los años 1950. El borrador del diseño de la red nacional de ordenadores incluía 100-200 grandes centros de datos en las principales ciudades, conectados a 20.000 centros de datos menores situados en agencias gubernamentales y grandes empresas.

Los grandes centros de datos estarían conectados mediante canales de banda ancha y abastecerían a diversos bancos de éstos encargados del procesamiento de datos, a los que cualquiera podría acceder desde cualquier terminal de la red, e

incluso eliminaría el papel moneda, al funcionar con transacciones electrónicas, recuperando el ideal marxista de una sociedad comunista sin dinero.

Paralelamente, en Estados Unidos, el gobierno estadounidense financiaba, a través de la Agencia de Proyectos de Investigación Avanzados de Defensa, un programa de investigación secreto para conseguir que el ejército estadounidense pudiera comunicarse a través de una red simple pero fiable, capaz de sobrevivir a catástrofes, y que no tuviera un punto de control central: internet. No se parecía en nada a los diseños de red existentes en aquella época, gestionados por empresas de telecomunicaciones y organismos de estandarización, que se basaban en estructuras de gestión muy rígida, como la que se planteaban los soviéticos.

Internet nació como una red de redes, en la que la estrecha coordinación entre las diferentes partes y los algoritmos de cooperación facilitaban la conexión de millones de nodulos sin tener que depender de una autoridad central de planificación.

Cuando murió Stalin, en 1953, la economía soviética basada en su colectivización forzada sufría graves carencias. El sistema de planificación central tenía dificultades para asignar cuotas de producción y la toma de decisiones jerárquica no proporcionaba incentivos a la iniciativa y la innovación; sencillamente no alcanzaba a cubrir la gran magnitud de aquella tarea, en una sociedad cada vez más compleja. En noviembre de 1962, Jruschov sugirió a sus camaradas del partido que tomaran como ejemplo la computación «racional» occidental y sus técnicas de gestión para solucionar los problemas del país.

Mientras tanto, y según los informes de la CIA, en octubre de 1962 el secretario de Kennedy escribió en una nota interna que el «compromiso total de los soviéticos con la cibernética» le daría a la Unión Soviética «una tremenda ventaja» y que tenían posibilidades reales de crear una auténtica internet. Otros proyectos soviéticos de tecnología a gran escala, como las armas nucleares y el programa espacial, habían conseguido superar desafíos tecnológicos mucho más graves, así que los estadounidenses estaban preocupados.

Sin embargo, los grandiosos planes de los técnicos cibernéticos de la Unión Soviética para la creación de una red nacional de centros informáticos con los que alcanzar una planificación y una gestión óptimas de la economía nacional no llegó a dar frutos, y los soviéticos se quedaron sin internet. El problema era que el proyecto de los soviéticos buscaba crear un sistema que se autorregulara, pero paradójicamente insistían en construirlo y liderarlo desde lo alto, por decreto ley... y fracasó, del mismo modo que estaban fallando los algoritmos del fútbol rígidos y con una planificación y liderazgo centralizados.

El fútbol como algoritmo de datos

En el Dinamo, Lobanovski reclutó a un joven profesor universitario, Anatoly Zelentsov, para que dirigiera el «laboratorio» del club, y para que evaluara la disposición funcional de los jugadores y el modo de desplegar al máximo su potencial por medio del análisis de datos. Era pseudociencia, pero demostraba mucha visión. Aun así, el ordenador central planificó un modo de juego que acabó dañando el fútbol ucraniano, puesto que los jugadores se empeñaban en completar las acciones planificadas por el ordenador, independientemente de su eficacia.

Estos planteamientos, con una planificación rígida y basados en decisiones informáticas, no proporcionaban los incentivos necesarios para que los jugadores interactuaran y se sintieran satisfechos, ya que la innovación del equipo quedaba anulada. Ésos fueron precisamente los motivos por los que la iniciativa soviética para crear una internet no dio fruto.

Con el tiempo, los algoritmos del fútbol y su liderazgo evolucionaron, pasando de un modelo planificado en dos fases —en el que se defendía o se atacaba— y con formaciones rígidas, al *tiki-taka*, que permitía que los jugadores adoptaran cualquier papel y que se dejaran llevar por una dinámica en la que lo importante era conservar la posesión y recuperar rápidamente el balón, y «bailar» por el campo creando oportunidades para marcar.

Internet, como el fútbol, es una red muy compleja y de éxito, pero una red de ordenadores. Lobanovski fue el primero en plantearse el fútbol como un algoritmo de red: una red de jugadores. Los objetivos de internet y del fútbol tienen, en cierto modo, se parecen. Internet necesita enviar información de un punto del mundo a otro punto del mundo lo más rápidamente posible, sin perderla. Del mismo modo, los jugadores de fútbol necesitan pasar el balón de una parte del campo hasta la meta del equipo rival, lo más rápida y eficientemente posible, sin perderlo.

En cierta medida se observa un paralelismo entre los algoritmos de la táctica del fútbol e internet, en el modo en que han evolucionado y se han transformado. Si pensamos en términos de estrategias, al principio los *routers* de internet a menudo se congestionaban —como el centro del campo, cuando se jugaba al fútbol con formaciones tradicionales— y había que redirigir la información siguiendo vías alternativas. Así, el redireccionamiento de internet evolucionó de sencillos algoritmos de direccionamiento que hacían pasar la información a través de una serie predeterminada de nodulos a unos algoritmos de direccionamiento altamente sofisticados basados en datos y en la IA que evitan la congestión de un modo dinámico, aunque tengan que pasar por múltiples *routers* que cooperen entre sí para hacer llegar la información al otro extremo de un modo rápido y fiable (igual que hicieron los jugadores del Kiev en su segundo gol contra el Atlético, haciendo pasar el balón por múltiples jugadores que tenían la posibilidad chutar pero que siguieron pasándolo).

En términos de formaciones, al principio el fútbol se basaba en formaciones en las que el balón llegaba rápidamente a los delanteros con uno o dos disparos largos;

luego se añadieron formaciones jerárquicas 2-3-5, para jugar con dos-tres pases, que producían muchos goles pero que eran muy pobres defensivamente, hasta llegar a la revolución del *tiki-taka*, en el que las posiciones desaparecían por completo.

Los primeros algoritmos de diseño de internet también eran de liderazgo jerárquico, con pocas conexiones intermedias: el ordenador del cliente conectaba con los Proveedores de Servicios de Internet (ISP), que a su vez conectaban con ISP regionales en un nivel más alto de la jerarquía, y éstos a su vez conectaban con ISP nacionales o internacionales, en lo más alto de la jerarquía, para llegar al contenido solicitado (por ejemplo, una página web o un vídeo).

Cada vez es más habitual que los medios públicos analicen los partidos de fútbol utilizando registros y algoritmos. Un ejemplo es el creciente número de blogs de internet de acceso gratuito que proporcionan análisis detallados de los partidos, que utilizan técnicas de observación basadas en las emisiones de televisión de esos partidos, así como bases de datos de fútbol disponibles al público en internet, con lo que proporcionan un nuevo tipo de análisis de juego —basado en datos—, muy al estilo de lo que se hacía en Estados Unidos con el béisbol a finales de los años 1990.

Hasta hace poco, el análisis táctico del fútbol profesional se basaba en registros observacionales y en el entrenamiento para adquirir experiencia, descartando la mayor parte de la información contextual. Sin embargo para el análisis de las tácticas de equipo, es necesario disponer de datos detallados de diversas fuentes, como la habilidad técnica, el rendimiento fisiológico individual, los esquemas de posiciones, el estilo y la estructura de la red, entre otros, para poder representar los procesos complejos responsables del comportamiento táctico del equipo: el alma de los datos del equipo. Aún no sabemos gran cosa sobre cómo influyen todos estos factores a la conducta táctica en el fútbol de elite. En parte se debe a la falta de datos. No obstante, cada vez disponemos de más registros de juego detallados obtenidos con tecnologías de rastreo de nueva generación, además de datos de entrenamiento fisiológico recopilados con innovadores sensores en miniatura, con los que poder investigar y crear el alma de los datos de un equipo.

Eso, no obstante, nos lleva al problema opuesto: que disponer de una cantidad ingente de datos se convierta en un obstáculo en sí mismo, al carecer de líneas metodológicas y modelos teóricos para decidir los algoritmos tácticos en el deporte. Ahí es donde los macrodatos, el entrenamiento de las máquinas modernas y las tecnologías de IA pueden ayudarnos a afrontar el problema, contribuyendo a desarrollar un modelo teórico para la toma de decisiones tácticas, las estructuras de cooperación y de liderazgo en los deportes de equipo. Para que funcione, primero hay que tratar con los propios clubes de fútbol para superar considerables obstáculos administrativos y poder gestionar los datos y acceder a estas tecnologías. No obstante, la introducción de los macrodatos y de las tecnologías de IA es una necesidad irrefrenable. Se avecina una época muy interesante para el análisis del rendimiento de los deportes en equipo y el diseño de nuevas formas de estrategia y

liderazgo en grupo: cada vez obtenemos más información táctica a partir del alma de los datos de cada equipo, y eso va a permitirnos obtener resultados más precisos y efectivos.

Hasta ahora, las estrategias y los diseños de redes de internet más avanzados se pueden atribuir a las redes de pares (P2P, *peer-to-peer*). Las redes de pares se hicieron famosas por su capacidad para distribuir música y películas a velocidades de vértigo, al hacer que los ordenadores de los participantes cooperaran entre sí (Napster, e-mule, bit torrent y ahora parte de los grandes diseños de centros de datos en la nube, como el de Amazon, o los usados por Microsoft para combatir ciberataques). En un momento determinado, las redes de pares eran responsables de más de un 50 por ciento del tráfico total de internet. No tenían ninguna autoridad de liderazgo central encargada de la planificación ni una arquitectura rígida, sino que más bien eran formaciones autoorganizadas de tipo enjambre, parecidas a las que crean las bandadas de aves migratorias, los bancos de peces o los enjambres de insectos, que muestran una conducta colectiva de propiedades autoformantes. Grandes partes de internet no funcionarían bien hoy de no ser por los algoritmos de pares.

Los algoritmos de pares son sistemas con estructuras muy fluidas que se definen solas. Hoy en día las tecnologías de pares son uno de los pilares de internet: dan velocidad a los centros de datos, son la base de las criptomonedas (*bitcoin*, *blockchain*), mantienen los datos actualizados, protegen tu ordenador de los ataques fulminantes de virus y gusanos y te permiten disfrutar de contenidos de alta calidad. Se han convertido en parte esencial de internet, haciéndola más resistente, distribuida, rápida y adaptable. Sin los principios de las redes de pares, el diseño de internet estaría un par de décadas por detrás de su estado actual.

Curiosamente, el tiki-taka y el fútbol total guardan muchos parecidos con los algoritmos de pares de internet. Por ejemplo, en el fútbol total un jugador puede adoptar cualquier papel (defensor o atacante), y con las redes de pares cualquier ordenador puede solicitar o distribuir contenidos. Además, al igual que en el *tiki-taka*, las redes de pares también se han alejado de las formaciones tradicionales jerárquicas y de liderazgo y gestión centralizada, volviéndose más orgánicas, flexibles y adaptables, sin una clara vocación de adquirir una estructura fija.

Los algoritmos por pares funcionan con unos principios muy simples y dependientes del contexto local que, bien aplicados, hacen que el sistema general también dé un gran rendimiento: a) para empezar un ordenador selecciona un pequeño subconjunto de ordenadores al azar con los que cooperar; b) los ordenadores priorizan la distribución de los fragmentos de un archivo —por ejemplo, de una película— menos representados entre sus pares (para evitar quedarse pendiente de una porción del archivo, cuando quizá ya se tenga el 90 por ciento del archivo descargado), y c) los ordenadores siguen tratando con otros ordenadores que les han

correspondido (por ejemplo, asociándose con otros ordenadores que les han dado tanto como lo que ellos les han enviado).

Esta última regla algorítmica se llama *tit-for-tat* y ha sido objeto de amplios estudios por parte de los científicos como sencillo modo teórico de intercambio para hacer que los sistemas cooperen de forma más eficiente. Para crear esas reglas, los científicos realizaron estudios avanzados de computación distribuida y tardaron más de una década en conseguir que funcionaran. Esto es importante, porque internet — como el fútbol— es un sistema muy complejo, y poner en práctica diseños algorítmicos que funcionen bien con reglas muy simples es fundamental para poder redimensionarlos después y trabajar con millones de ordenadores, o aplicando estrategias de juego muy complejas en el campo de fútbol.

Desde la llegada del ordenador personal, unos cuantos pioneros del fútbol han usado los ordenadores y el análisis de datos con diversos motivos. El amigo de Lobanovski, el profesor Anatoly Zelentsov, desarrolló juegos de ordenador para poner a prueba a los jugadores. No obstante el gran paso adelante se dio en 1996, cuando la empresa Opta Index empezó a recoger «datos de partidos» de la primera división inglesa.

Hoy en día los atletas no sólo están controlados por las cámaras en los estadios, sino también por curiosos artefactos, como los acelerómetros, los pulsómetros e incluso sistemas parecidos a un GPS de corto rango. Desde luego, los macrodatos, la IA y las tecnologías modernas están revolucionando el deporte y creando una representación del alma de los datos de cada jugador y de cada equipo.

Los clubs deportivos, los medios de comunicación y los aficionados de todo el mundo comparten una pasión común por las estadísticas y la información. Los grandes clubs las usan para mejorar el rendimiento de sus propios jugadores, para preparar tácticas contra otros equipos o seleccionar jugadores potencialmente interesantes. Por otra parte, a los medios les encantan los datos, ya que dan un valor añadido a sus artículos. Por último, las estadísticas también han dejado en un segundo plano a los directores deportivos fantasiosos que crean sus equipos de fantasía fichando deportistas en su mejor forma, uno a uno.

Durante el reciente campeonato mundial de Brasil, los alemanes, famosos por sus conocimientos técnicos, tenían un as en la manga. Muchos aficionados al fútbol se mostraron escépticos cuando se reveló que la selección nacional llevaba puesto el sistema miCoach de Adidas durante las sesiones de entrenamiento, antes y durante la competición. Este servicio de monitorización fisiológica recoge datos y los transmite directamente desde el cuerpo del atleta, entre ellos la frecuencia cardíaca, la distancia, la velocidad, la aceleración y la potencia, y muestra esos valores al momento en un iPad, para que el equipo técnico puedan verlos durante el entrenamiento, así como en una sesión posterior de análisis en profundidad. Curiosamente, su análisis puede ayudar a identificar a los jugadores más en forma, frente a los que quizá necesiten un descanso.

La empresa Oulala Games Ltd, con sede en Malta, construyó una matriz matemática que usa los datos de fútbol de Opta Index para crear un eficiente sistema de puntuación. La plataforma creada por la empresa utiliza un sofisticado algoritmo para determinar los aspectos cruciales del rendimiento del deportista que contribuyen al resultado general. Su sistema incluye un total de 70 criterios diferentes, dependiendo de la posición del jugador (portero, defensor, centrocampista o delantero), que generan un total de 275 modos de sumar o restar puntos. Los valores combinados de estas acciones, correspondientes a jugadores incluidos en equipos simulados por un ordenador, dan como resultado los ganadores de las diferentes ligas diarias.

Da la impresión de que el fútbol, que siempre ha sido un juego de números, cada vez depende más de los macrodatos y de la IA. Muchos clubes dejan de contratar a cazatalentos y en su lugar fichan a analistas informáticos; la televisión, la radio y los periódicos hablan más que nunca de datos estadísticos sobre la actuación de los jugadores, de los entrenadores y de los equipos. Los números del fútbol también están invadiendo el campo de la fantasía: las estadísticas que rodean a los jugadores no sólo se usan para medir su rendimiento real, sino también para crear grandes equipos de fútbol o para mejorar las apuestas. De hecho, con menos de 10.000 líneas de código y muchos datos, es fácil construir un sistema de predicciones para las apuestas tan fiable como el mejor, sin tener que recurrir a montones de expertos o modelos complejos. Podemos decir que esta revolución de los macrodatos y la IA en el fútbol va a seguir adelante, y que cambiará la experiencia del espectador del deporte más popular del mundo.

En 2003, la revolución de los datos en el fútbol recibió un nuevo impulso desde el otro lado del Atlántico, cuando Michael Lewis publicó *Moneyball*, obra pionera que utilizaba nuevos datos estadísticos para valorar los jugadores de béisbol. En los últimos años el baloncesto —deporte muy fluido— ha encontrado excelentes aplicaciones para éstos, y la nueva era de recogida y procesamiento de macrodatos y de la IA ha facilitado las comparaciones entre jugadores y entre clubes, quién pasa el balón a quién, a qué velocidad corren los jugadores, quién tiende a iniciar los ataques más peligrosos del equipo, etc.

Sin embargo, estas estadísticas son útiles sobre todo para los deportes en los que el rendimiento individual del jugador es la clave del éxito. No obstante, en el fútbol, tal como había comprendido perfectamente Lobanovski, no eran tan importantes los jugadores individuales como la red de jugadores, y el liderazgo del equipo: ¡el alma de los datos del equipo como unidad!

Un laboratorio de datos en el campo

El profesor Paco Seirulo era el director del «laboratorio» del FC Barcelona, del entrenamiento físico y la planificación del equipo de Guardiola. Era el equivalente al profesor Zelentsov de Lobanovski. El profesor Seirulo estudió química, luego arquitectura, y por fin se hizo entrenador físico. Hace unos años, en Barcelona, tuve la suerte de conocer a Paco Seirulo, un hombre que según muchos, ha sido en gran parte responsable del desarrollo de los algoritmos del estilo *tiki-taka* y de la filosofía detrás la estrategia del FC Barcelona. Imagínate la escena: yo, un científico informático, sentado junto a uno de los entrenadores más respetados del fútbol moderno. Yo me he pasado la vida trabajando en lugares como Microsoft, Bell Labs o Telefónica; él se ha pasado la vida con futbolistas, boxeadores, nadadores y toreros (¡sí, también ha entrenado a toreros!).

Podría haber sido como una escena cómica de *The Big Bang Theory*, Sheldon encontrándose con Alex Ferguson: nada en común, un montón de silencios incómodos, ambos pensando: «¡Este tipo es de otro planeta!». Pero no, nada que ver. Lo más sorprendente era que ambos decíamos cosas similares, sólo que con un vocabulario diferente. Veíamos el mundo de acuerdo con el mismo esquema general. Cuando él hablaba del balón, yo pensaba en redes de datos y algoritmos. Cuando él hablaba de jugadores, yo pensaba en nodulos. Cuando él hablaba de pases, yo pensaba en una matriz; cuando él hablaba de la longitud y de la frecuencia de los pases, yo pensaba en pesos y frecuencias.

Así pues, mientras él me explicaba la evolución de la táctica del fútbol desde Lobanovski a Michels y al Barça de Guardiola (a cuya creación contribuyó él mismo), yo veía tantos paralelismos con el modo en que se habían desarrollado los algoritmos, pasando de una estructura de liderazgo jerárquica a un sistema de pares. Internet, al principio, era como el fútbol de Lobanovski, con poca cooperación entre los nodulos, pero el tiki-taka para mí era como los algoritmos de pares. Yo no dejaba de preguntarme si, del mismo modo que la estructura de pares estaba gobernada por sencillos algoritmos científicos, también lo estaba el tiki-taka, si habría reglas científicas básicas que lo gobernarán.

Entonces él me contó su problema. Al igual que en las estructuras de pares, los jugadores también usaban una serie de sencillas reglas para hacer posible un sofisticado sistema de cooperación en el campo, como con los algoritmos del tiki-taka. No obstante, todas estas estrategias y dinámicas se basaban en la intuición de esos grandes entrenadores, y en numerosos experimentos realizados en el campo y en los entrenamientos. Seguían basándose en su intuición, pero carecían de las herramientas científicas para comprobar cuantitativamente su teoría en los ordenadores, y encontrar una explicación. El conjunto de reglas que conocían no podía introducirse en un ordenador y analizarse, y mejorarse. Y ahí es donde vi que los macrodatos y las herramientas de extracción de patrones y aprendizaje automático que los expertos informáticos como yo habían desarrollado para internet también podrían aplicarse a los diferentes estilos de fútbol, y al alma de las tácticas.

Se ha demostrado que intentar definir la naturaleza del juego en el fútbol a partir de los datos reunidos en el campo de juego es un negocio frustrante para los científicos del deporte. La mayoría de estudios estadísticos revelan lo que los entrenadores profesionales ya ven con sus propios ojos.

No obstante, la teoría de algoritmos, la recogida de datos detallados y los últimos avances en computación de estos últimos pueden revelar unas complejas interacciones que se producen en el campo entre los jugadores que rodean el balón y más allá, hasta donde no llega el ojo humano. De hecho, hay muchos fenómenos de segundo orden en la interacción de los jugadores que tienen un enorme impacto en la dinámica del juego y que no se pueden observar a simple vista, pero en los que pueden ser de gran ayuda los datos y los análisis por ordenador.

Una primera demostración de las ventajas de usar las técnicas científicas de redes y datos fue la que dio un grupo de científicos del UCL (University College of London) y la Queen Mary University durante la final de la Eurocopa de 2012. Usaron el famoso algoritmo Page Rank de Google (que ayuda a crear rankings de páginas web) para crear un ranking de importancia de los jugadores en un partido basándose en el número de pases que reciben o qué jugadores son cruciales para mantener las conexiones del equipo. Eso da una idea aproximada de quién es más probable que acabe teniendo el balón tras una serie de pases. En el caso del Barça, era Xavi, considerado uno de los más grandes centrocampistas de todos los tiempos.

No obstante, su trabajo, como otros, se centraba en estadísticas de alto nivel de la red de pases o en la consistencia de la red de jugadores. Eran mediciones de valor acumulativo a lo largo de diversos partidos, por lo que no podían capturar la dinámica de un partido específico, o la evolución de la red con el tiempo, el orden de los pases o el grupo de jugadores por los que pasa el balón... Básicamente, los flujos de algoritmos, los motivos o los «diseños» del juego.

Precisamente es la red de pases y la posición y la dinámica de algoritmos de los jugadores en el campo lo que ha fascinado a científicos especializados en redes como yo. La estrategia del pase se asocia con el éxito. La red de pases de un equipo de fútbol se compone de los jugadores, como vértices, y de los pases entre jugadores, como aristas. Eso crea una red como internet, que puede estudiarse por sus características únicas. Analizando el algoritmo dinámico entre jugadores de fútbol, se puede determinar qué es lo que hace que algunos equipos de fútbol jueguen un tipo de fútbol completamente diferente al de otros.

En la última década, diversas empresas proveedoras de datos y sitios web han creado un registro de éstos de los partidos de fútbol, y publican series de datos con estadísticas sobre lo que ocurre en cada partido y lo que hace cada jugador, lo que ofrece un alma de los datos del campo. Entre estas iniciativas se cuentan Prozone, OptaPro, Instant Football y Squawka, por ejemplo. Hoy en día los estadios están equipados con cámaras de alta definición que se pueden usar para crear un mapa de posiciones de cada jugador y del balón en el campo a cada décima de segundo, con

una resolución de 10 centímetros, y durante los entrenamientos y los partidos de preparación se recogen referencias GPS fijando a los jugadores mecanismos especializados para la recogida de datos. La prevalencia de éstos es lo que nos permite buscar el alma de los datos del equipo para decodificar su particular estilo de juego y cómo contribuye a él cada jugador.

El último gran éxito del análisis de datos en el fútbol es sin duda el equipo británico del Leicester, que consiguió ganar la liga inglesa después de que al empezar el campeonato las apuestas le dieran una probabilidad de 1 contra 5.000. Uno de sus muchos secretos fue el uso de equipamiento Prozone durante las sesiones de entrenamiento y el procesamiento de los datos con la herramienta analítica Prozone 3. Eso hizo posible personalizar las sesiones de entrenamiento para sacar lo mejor de cada futbolista individualmente y reduciendo al mismo tiempo el riesgo de lesiones. Es un dato importante, ya que la mayoría de los especialistas en fútbol basan el éxito del Leicester en el reducido número de lesiones sufrido por el equipo durante esa temporada. El equipo también recurrió a la suite de productos analíticos desarrollados por OptaPro, solución que proporciona informes interactivos que se pueden ver en un iPad antes, después y durante un partido. En su feudo, el King Power Stadium, el Leicester tiene una «sala de guerra» para sus analistas de datos, conectada con el vestuario, para que los jugadores y el personal técnico dispongan de las estadísticas durante la media parte. Hoy en día aún es difícil evaluar qué beneficio real aportará el uso intensivo de su recogida y la IA. Pero el éxito en muchos casos depende de cosas muy pequeñas. ¡Así que cualquier ventaja, por pequeña que sea, vale la pena!

En el análisis deportivo, cada vez tiene más importancia la ciencia de redes algorítmicas, que tratan a cada jugador como un nódulo y trazan una línea entre ellos cuando el balón viaja de uno a otro. Éste ha sido un campo de investigación que ha dado frutos porque ya se han podido desarrollar una serie de herramientas matemáticas con datos para analizar las redes. También se puede considerar el fútbol como un juego de invasión. Estos juegos consisten en dos equipos que compiten por la posesión del balón en una zona de juego limitada. Ambos equipos tienen el objetivo de marcar llevando el balón a la meta contraria, y también de defender su propia meta. El equipo que marca el mayor número de goles al final del partido es el ganador del encuentro.

Los jugadores de fútbol construyen redes en el campo con sus pases y con el patrón que crean esos pases con el tiempo, por ejemplo, el conjunto de jugadores implicados, el orden de los jugadores, etc. A eso se le llama los «diseños de pases». Hay diseños de pases simples y otros más complejos. El algoritmo de los diseños de pases simples tiende a crear redes simples fáciles de superar, y no crea muchas ocasiones de gol, mientras que los diseños de pases más ricos crean redes más complejas y una dinámica de juego con mayores oportunidades.

Por ejemplo, el científico Laszlo Gyarmati demostró que en el fútbol moderno se repiten cinco patrones diferentes de secuencias de tres pases: ABAB, ABAC, ABCA,

ABCB, y ABCD. La secuencia ABAB, por ejemplo, es cuando el jugador 1 pasa el balón al jugador 2, que se lo devuelve al jugador 1, y éste vuelve a pasárselo al jugador 2. Si participan dos jugadores más, el 3 y el 4, sigue siendo una secuencia ABAB. Secuencias como la ABCD crean redes simples, pero otras como la ABCB crean redes más complejas.

Las innovaciones son frecuentes, y los equipos no dejan de adoptar o abandonar nuevas tácticas, por lo que es fácil suponer que difícilmente se podrá jugar un único tipo de fútbol, con tácticas y habilidades que no vaya a copiar ningún otro equipo. Pero eso no es cierto del todo. De hecho, usando técnicas de análisis de datos mi equipo ha podido determinar que, aunque muchos equipos compartan estilos de juego similares, hay uno que destaca por su personalidad única, que juega un tipo de fútbol que ningún otro equipo puede igualar, y que sigue unas reglas únicas. Ese equipo es el FC Barcelona de Guardiola, uno de los equipos de fútbol más exitosos del planeta.

Los datos y la IA revelan patrones particulares y reglas simples que resultan difíciles de definir usando elementos como los goles, las faltas, los saques de esquina, etc.

Así pues, ¿hasta qué punto es único el *tiki-taka*? Bueno, cuando representamos gráficamente las secuencias desplegadas por equipos muy diferentes, resulta que el Barcelona destaca claramente por usar series ABCB y ABAC (por ejemplo, Xavi a Messi, de vuelta a Xavi, y éste a Neymar) con una frecuencia mucho mayor, mientras que recurre menos a las secuencias ABCA y ABCD. Ése es un dato fascinante sobre la naturaleza del fútbol. Además, a diferencia de otros valores usados para medir los deportes, da una clara indicación de en qué se diferencia el juego del Barcelona del de los demás. En otras palabras, el juego del Barcelona tiene mucha más estructura y unos algoritmos más bonitos que la simple definición del *tiki-taka* como una serie indefinida de pases al azar, en lo que es un sistema cuidadosamente construido. El artículo «FC Barcelona, The Science behind their success», publicado por la BBC, afirma que el estilo de juego del Barça, el *tiki-taka*, ha revolucionado el fútbol, y que este análisis revela que el equipo que muchos consideran el mejor de la última década tiene un estilo propio.

También puede usarse un tipo de análisis parecido para encontrar nuevos jugadores por el mundo e integrarlos al equipo, para identificar nuevos talentos que encajen en el estilo de juego del Barça. Por ejemplo se puede analizar las secuencias y los patrones de pase de un jugador y su movimiento en el campo antes y después del pase para categorizarlo. Si por ejemplo hubiera que sustituir a Gerard Piqué (el famoso defensa central del FC Barcelona) con algún otro jugador similar, usando los macrodatos y el análisis con IA de las secuencias pasadas de cientos de futbolistas, concluiríamos que Jeremy Mathieu (Valencia y después FC Barcelona), Képler Laverán Lima Ferreira —Pepe— (Real Madrid y luego Besiktas), Javier Mascherano

(Liverpool y luego FC Barcelona), Gustavo Cabral (Celta de Vigo) y Nacho Fernández Iglesias (Real Madrid) muestran unos diseños de pase similares y probablemente serían buenos sustitutos. No es de extrañar que dos de esos jugadores hayan pasado por el FC Barcelona.

Para aplicar el *tiki-taka* es necesario que todos los jugadores del equipo estén sincronizados y muy preparados para jugar de ese modo tan armónico, y eso es difícil. Si los jugadores juegan ligeramente peor de lo que deberían, como equipo y como red, les irá mucho peor, y los errores en el fútbol pueden salir muy caros. Los algoritmos del *tiki-taka* ayudan a conservar la posesión del balón, pero no aseguran las oportunidades de marcar, y requieren complejas interacciones entre los jugadores alejados de la pelota. Así pues, el *tiki-taka* funciona bien si se aplica muy bien, pero también se puede tener un 90 por ciento de la posesión y perder el partido.

Recientemente, equipos como el Real Madrid o el Bayern de Múnich han empezado a ganar al Barcelona. De hecho, Mourinho, el entrenador del Real Madrid, afirmaba que no se trata de conservar el balón, como parece que hace el *tiki-taka*, sino de beneficiarse de un error del oponente con jugadores rápidos y precisos que no pierdan una oportunidad. Un error puede tener consecuencias catastróficas, aseguraba.

Guardiola dejó el Barcelona para entrenar al Bayern de Múnich, que en 2013 se clasificó para la final de la Champions League superando al Barcelona por un resultado global de 7 a 0. Luego la selección nacional alemana ganó el Mundial de 2014. Hace unos meses se dijo que Guardiola había declarado que «odiaba el *tiki-taka*» después de haberse pasado, según decían, a un sistema más equilibrado que pone el énfasis en unas transiciones más rápidas y en un juego más vertical. Guardiola criticó el *tiki-taka* y empezó a jugar con un posicionamiento 3-4-3, el esquema táctico simétrico más vertical del fútbol. Volvió a las tácticas que le habían dado fama y reconocimiento: un fútbol de posesión llevado casi al límite. Hubo quien llegó a representar el esquema de Guardiola con el portero al borde del círculo central y todos sus defensores al menos 15 metros dentro del campo contrario.

Con Guardiola, el Bayern pasó a convertirse en un equipo mucho más impredecible. Ahora pueden jugar de diferentes modos, característica propia de los grandes equipos. Hay más interacciones en vertical, y un mayor control, gracias a la acción de los mediocampistas. Todo ello supone una progresión positiva. Hace poco se publicó que Guardiola había dicho a los suyos: «Sed vosotros mismos. Tenéis que buscar en vuestro ADN. Odio el *tiki-taka* como concepto simple. El *tiki-taka* significa pasarse el balón sin más, sin una intención clara. Y no tiene sentido. El secreto es sobrecargar un lado del campo para que la defensa del contrario se vea obligada a bascular, dejando el otro lado en situación vulnerable. Y cuando hayamos hecho eso, tenemos que atacar y marcar desde el otro lado. Por eso tenéis que pasar el balón».

Pero ¿cuál es el algoritmo correcto para una estrategia que sobrecargue y debilite la defensa del equipo contrario? ¿Existe sólo una? En el campo de la investigación de redes, los últimos diseños de redes de la internet del futuro y la evolución de los sistemas de pares hacen pensar que quizá lo que mejor funcione sean unos diseños más híbridos entre sistemas posicionales fijos y de distribución cooperativa, unos diseños con una combinación de sistemas de gestión centralizada y otros más orgánicos y de organización independiente.

¿Podría trasladarse también esta evolución al fútbol? ¿Y si la próxima filosofía de la estrategia del fútbol se basa no ya en ocupar de un modo fluido el lado del campo que más convenga según surja la ocasión, sino también en reducir al mínimo el tiempo en que se crea una oportunidad de marcar con un juego más vertical? ¿Y si hace falta una combinación de formaciones jerárquicas y sistemas de pares, al estilo tiki-taka, para especializar determinadas posiciones y obtener mayores velocidades de transición? ¿Y si resulta que hay que complementar los enfoques como el tiki-taka, que son muy resistentes y reactivos —lo que se traduce, por ejemplo, en que los jugadores casi nunca pierden el balón y en que tienden a recuperarlo rápidamente— con arquitecturas más jerárquicas que permitan mejorar la eficiencia del juego y que requieran menos coordinación entre los jugadores? ¿Y si hay que sacrificar un poco el juego vistoso para conseguir una defensa más robusta o una mayor velocidad de ataque? ¿Y si al final no existe una única estrategia mejor? ¿Y si la mejor estrategia es contar con diversas estrategias, aprender y usar una combinación de estrategias a medida que progresa el juego? ¿Y si todo esto encontrara respuesta a través del alma de los datos de un equipo y la IA que aprende de forma continua?

Sería fácil repasar los últimos años del fútbol y concluir que el juego ha alcanzado su madurez. No obstante, el fútbol y sus técnicas no dejan de cambiar y de adaptarse con el tiempo. Es difícil de creer que las mejoras en el análisis detallado de datos con ordenadores y la IA no vayan a traer un gran adelanto.

Cuanto más sofisticadas sean las tecnologías, más sofisticados serán los sistemas. El aprendizaje automático, la IA y el estudio de los datos pueden llevar la estrategia del fútbol a un nivel de comprensión y de innovación mucho más profundo, centrándose en las dinámicas entre jugadores y analizando los patrones que emergen. Tenemos que pasar de una estadística estática reduccionista a una nueva métrica de datos que capture la dinámica algorítmica y las estructuras de interacción entre jugadores y sus pases en el campo, y que aprenda y evolucione de forma automática para modificar lo que sucede durante el entrenamiento y en los partidos.

El futuro de la preparación de los partidos es la IA. Usando los macrodatos y la IA con cada jugador, se pueden usar modelos individuales de comportamiento detallado para cada juego, usando un aprendizaje de imitación profunda para predecir la respuesta más probable de un jugador a una situación de juego propuesta y plasmar de un modo visual la respuesta más probable. El futuro de la investigación del deporte pasa por un enfoque multidisciplinario más fuerte. Será necesaria la

colaboración de analistas de rendimiento, expertos del ejercicio, biomecánicos y deportistas para interpretar estos complejos conjuntos de datos.

Por ejemplo, la mayoría de iniciativas de aplicación del aprendizaje automático al fútbol son obra de grupos de investigación científica en ordenadores, cuya presencia no es habitual en la mayoría de clubes de fútbol ahora mismo. También se observa que quizá sean las futuras colaboraciones entre científicos informáticos y del deporte las que consigan aplicar estos enfoques complejos de una forma más relevante. A su vez, confiar cada vez más en complejas técnicas de análisis de datos puede plantear nuevos desafíos a los futuros científicos del deporte.

Cruyff solía criticar el uso de estadísticas obtenidas por ordenador como argumento para rechazar a un jugador. Alegaba que, si se hubiera basado en ese criterio, el Ajax le habría rechazado a él. Estaba seguro de que cuando tenía quince años su técnica y su visión no habrían sido detectables para un ordenador. Yo quería creer que en unos años y con las técnicas de aprendizaje automático existentes, con los nuevos tipos de datos que se recogen y el poder de los ordenadores para descubrir nuevas señales sutiles que escapan a la observación a simple vista, Johan Cruyff habría destacado claramente.

Han sido necesarios años, si no ya décadas, para perfeccionar el tiki-taka, pero con los innovadores análisis de datos y modelos matemáticos que simulan el juego, puede que no tengamos que esperar tanto para la aparición de un nuevo estilo revolucionario de juego, y la próxima revolución quizá no surja de un entrenador, como fue el caso de Lobanovski o Guardiola, sino de un cerebritito informático o un científico experto en IA.

Podríamos pensar que van a aparecer complejos sistemas de IA que capturen y simulen el comportamiento de un equipo frente a otro recurriendo a una gran cantidad de análisis de datos y que intente predecir el resultado (por ejemplo, posesión, recuperaciones, distribución espacial de los jugadores, patrones de pases, marcadores, etc.) y el *ranking* de los diversos equipos. Podríamos analizar las acciones tácticas que tienen que jugar en el campo y el rendimiento, no sólo en términos de número de pases, sino también en cómo influye el jugador A en el comportamiento del jugador B, por ejemplo si atraen jugadores a una determinada zona al pasarse el balón, sobrecargando esa zona para descargar otra, y creando oportunidades para marcar.

Hoy en día, los equipos de fútbol modernos están empezando a crear laboratorios para estudiar los datos y aplicar las lecciones de Lobanovski y las nuevas dinámicas del juego. Todos los partidos del equipo son analizados en todas sus dimensiones posibles por científicos expertos en datos deportivos que hacen uso de los macrodatos. En el pasado, para entrenar era fundamental la visión y la experiencia. Los grandes entrenadores de la historia del fútbol —de Helenio Herrera a Arrigo Sacchi o Pep Guardiola— eran capaces de sacar grandes conclusiones de los patrones del fútbol y de aplicarlas a nuevas y revolucionarias estrategias de éxito.

La situación está cambiando. Además de un equipo de jugadores, los clubs de fútbol están fichando nuevos equipos de analistas de datos deportivos y IA con el objetivo de extraer datos que permitan hallar nuevas tácticas y patrones algorítmicos ganadores. Los equipos que usan la ciencia de los datos para mejorar el rendimiento del equipo y de los jugadores, o para buscar nuevos jugadores obtendrán una enorme ventaja con respecto a sus competidores. Pero ¿tiene límites el uso de los macrodatos y IA en encontrar nuevas formas de liderazgo y cooperación, en reinventar el deporte de equipo?

Alan Turing, considerado la figura más destacada de la ciencia informática, con trabajos decisivos en teoría matemática, filosofía, IA e ingeniería práctica, desarrolló la máquina Enigma, que ayudó a descodificar los mensajes alemanes durante la guerra. Winston Churchill declaró que Turing había hecho la mayor contribución a la victoria aliada en la guerra contra la Alemania nazi al permitir descifrar los mensajes en código interceptados, lo que permitió derrotar a los nazis en diversas batallas cruciales. El héroe de la segunda guerra mundial era un informático.

No obstante, uno de los resultados más determinantes de Turing fue que demostró las limitaciones y los problemas de los ordenadores. Se podría pensar que los ordenadores, los datos y los algoritmos complejos tienen la capacidad de resolver hasta el más complicado de los problemas y de las estrategias, no sólo del deporte y el liderazgo, sino también de cualquier otra disciplina, que no hay problema que un ordenador no pueda resolver. Pero Turing demostró que la respuesta es un no rotundo, que hay problemas que nunca resolverán los ordenadores, la IA ni los macrodatos.

El hecho es que hay cosas que son computables y otras que no, y eso plantea un interesante contrapunto a los numerosos algoritmos y a las revoluciones informáticas y de IA que podemos esperar, el saber que siempre habrá problemas cuyas respuestas no son computables. Es aquí donde entra en juego la magia, la intuición, el genio de los entrenadores, de los jugadores, de los seguidores, y su pasión, y eso no desaparecerá nunca. Es lo que hace que el fútbol sea un juego tan mágico, en el que una esfera perfecta entra en contacto con el burdo pie de los jugadores para entretenimiento de tantos. Es aquí donde aparecerán una y otra vez nuevos «cisnes negros», eventos que proporcionen ventajas que los ordenadores no consigan predecir, porque usar el aprendizaje automático y la IA es como usar el pasado para predecir el futuro, y a veces el futuro no es predecible.

Muchos escritores, periodistas e incluso directivos nos dicen que hemos asistido a la última gran revolución táctica del fútbol. Muchos creen que hemos llegado al final del camino en lo relativo a táctica y el liderazgo de equipo.

Roberto Mancini decía que los futuros avances en el fútbol vendrían no de la táctica, sino de la preparación física de los jugadores. Jonathan Wilson, reputado

redactor de deportes de *The Guardian*, afirma que el fútbol es un deporte maduro que ya sido examinado y analizado exhaustivamente durante casi siglo y medio y que, suponiendo que el número de jugadores no varíe, probablemente no quede ninguna pendiente ninguna revolución que vaya a asombrar al mundo. Yo no pienso así. No sé cuál será la próxima gran revelación táctica, y de liderazgo y colaboración de equipos, pero sí que creo que la habrá, y que se encontrará con los ordenadores, los macrodatos y la IA.

Los ordenadores nunca podrán sustituir a los ¿entrenadores?, o por lo menos no de momento. El juego del fútbol es uno de esos problemas que un ordenador no puede resolver por completo, pero los ordenadores pueden ayudar, y el alma de los datos de un equipo puede desvelar patrones y modos de jugar al fútbol nunca vistos hasta ahora. Algo parecido puede pasar en el mundo de la creatividad.

El ordenador en la cocina

En los años 1990 había muchas tareas que, indiscutiblemente, requerían la inteligencia de los humanos, y se les pagaba por su grado de experiencia y especialización: por ejemplo, para planificar un viaje en avión con varias paradas. En 1990 era esencial contar con un agente de viajes para encontrar un itinerario corto y de bajo coste. Hoy, en cambio, esta tarea la hacen mejor los ordenadores, capaces de analizar una cantidad de datos mucho mayor. Con la evolución de la tecnología, esta transformación gradual de las tareas, que pasaban de ser aparentemente intuitivas a evidentemente mecánicas (y por tanto aptas para un ordenador) sigue avanzando. Tanto si consideramos estas tareas como cuestión de «inteligencia» o no, es de esperar que en los años venideros haya cada vez más.

Por ejemplo, en los últimos veinte años IBM ha demostrado que un ordenador puede analizar todo tipo de combinaciones posibles en una partida de ajedrez y vencer al mejor jugador de ajedrez del mundo. Pero ¿y si un ordenador pudiera analizar todas las combinaciones de todos los ingredientes del mundo que usamos hoy en día y crear o sugerir otros nuevos con un sabor sorprendentemente bueno? En 2015 IBM desarrolló el Chef Watson, un ordenador que analiza el contenido de más de 9.000 recetas, sacando información sobre patrones de elaboración y terminología de esas combinaciones. Combinando esos datos con su gran conocimiento de la química de los alimentos y con las preferencias de sabores de los seres humanos, Watson era capaz de crear recetas totalmente nuevas con nuevas combinaciones de ingredientes.

Antes este tipo de labor creativa —elaborar nuevas recetas, descubrir nuevas combinaciones de sabores— era prerrogativa de los chefs humanos. Pero los expertos en computación están demostrando que los algoritmos de IA también pueden ser creativos.

Los ordenadores han sido capaces de combinar ingredientes de formas aleatorias durante muchos años, pero eso no significa que un bocadillo de chocolate, ketchup y coles de Bruselas sea una buena idea. Así pues, ¿cómo se va más allá de las combinaciones al azar para encontrar otras creativas?

De hecho, la ciencia interviene en gran medida. El ordenador debe saber, por ejemplo, qué proporciones de harina, mantequilla y levadura son las más frecuentes en las recetas de las galletas o los pasteles para crear una receta que produzca el resultado deseado.

Los expertos de IBM empezaron a «entrenar» a Watson —que se llama así en honor a Thomas Watson, fundador de la empresa— con decenas de miles de recetas, empezando por la composición química de los alimentos y pasando luego a los sabores e ingredientes que se complementan mejor —como el romero y las patatas, o las aceitunas y la ginebra—. Pero aunque el Chef Watson es capaz de comprender enormes cantidades de datos y de aprender de las interacciones con las personas y la información, de pasar de sacar conclusiones del trabajo de los chefs a generar buenos platos hay un buen trecho, y siempre requerirá el genio y la creatividad de los grandes chefs. Hace poco, un periodista que había usado el Chef Watson de IBM declaró que el ordenador había creado el peor burrito de la historia, con una combinación de ingredientes imposible, y se preguntaba cómo podía ser tan mala la receta. Así que aún queda mucho para que los ordenadores puedan crear platos al nivel de los humanos.

La química, la biología y la comprensión de la física de la comida ha influido en la última generación de chefs y en su exploración de nuevos conceptos y filosofías. Pero ¿cómo influirán en ellos los datos y la IA?

El cocreador de la gastronomía molecular, Hervé This, se ha convertido en el científico loco de la gastronomía más popular de Francia. Recientemente This ha empezado a trabajar en lo que él llama cocina nota a nota, en la que los ingredientes se descomponen en texturas, sabores y compuestos individuales altamente deconstruidos para comprender la fisiología del sabor y crear nuevas experiencias.

Por ejemplo, puede crear un postre consistente en gominolas de manzana, granizado de limón y gofres de caramelo, pero para el que no se usan ni manzanas, ni limones ni caramelo: los sabores proceden del ácido cítrico y de un compuesto de glucopiranosil-D-sorbitol, un sustituto del azúcar. O una crepe con sabor a filete, al que se le puede añadir un compuesto de metanol para que dé la impresión de que hay patatas en el interior del filete, y otra combinación de compuestos que pueda darle sabor de pizza o de palomitas... la lista de combinaciones es interminable. Esto no son más que pequeños pasos en el complejo proceso de la creatividad en cocina nota a nota, que se enseña en varias escuelas culinarias punteras en Europa, y que puede acelerarse con la identificación de nuevos compuestos y combinaciones, gracias a la ayuda de la ciencia de los macrodatos y la IA.

Con esta deconstrucción los alimentos pueden convertirse en espumas, geles y otras materias no inmediatamente reconocibles como comida. No obstante, la gente no se come lo que no reconoce, así que la presentación de todas estas nuevas combinaciones de sabores debe tener unas formas y elementos que recuerden platos conocidos u objetos familiares. Existe un concepto llamado neofobia alimentaria, que consiste en que la gente se niega a comer cosas extrañas, y ahí es donde entra la poesía y la fantasía de los chefs para ganarse al público.

Desde un punto de vista científico, basado en la química y la física, los sabores de los ingredientes (que son compuestos químicos) se consideran unidades mínimas de alimento que permiten una interpretación más básica de los diferentes ingredientes y de sus propiedades. De este modo, cada ingrediente puede ser discriminado por sus sabores, que son percibidos por el sistema sensorial humano (la nariz, la boca, etc.). Percibimos los alimentos a través de diversos procesos cognitivos, como el sabor, la imagen, el olor, la textura, etc.

Por ejemplo, el sabor se percibe a través de diversos receptores químicos de la boca y de la nariz, y se calcula que el 90 por ciento de la percepción del sabor depende de la discriminación de lo que llamamos sabores volátiles en los receptores de la nariz. También hay un conjunto finito de compuestos químicos que son específicos de diversos sabores percibidos por los seres humanos. No obstante, percibimos el sabor sobre todo por el olfato. Dedicamos el 2 por ciento de nuestro genoma a la capacidad de oler diferentes cosas (y ésta es prácticamente la misma cantidad de genes que nos separa de los chimpancés).

Los aderezos, en general, se componen de una cantidad mínima de sabores que les dan su sabor característico: por ejemplo el nabo (3 sabores), la salsa de soja (7 sabores), el tomillo (28 sabores), el romero (31 sabores), etc. Los aderezos suelen usarse para acentuar un sabor particular y tienen sabores característicos. En cambio, el tomate se cuenta entre los 10 ingredientes con mayor número de sabores del mundo (unos 164 sabores). Eso explicaría que pueda combinarse con tantos otros ingredientes y que pueda ser un ingrediente de base en recetas de tantas culturas y regiones. Otros ingredientes que tienen esta propiedad son la manzana, la cerveza, el cacao, el café, el coñac, el ron y el té.

Del mismo modo que los ingredientes presentan un número de combinaciones casi infinito, la combinación química de sabores también puede producir un número de posibilidades (de combinaciones) muy grande. La búsqueda creativa de combinaciones acertadas entre todas las posibilidades siempre ha sido —y sigue siendo— todo un reto. En esta búsqueda, a través del método de ensayo y error, de la experiencia y de la intuición, los chefs han ido creando series de afinidades entre ingredientes (ingredientes que combinan bien entre sí, por sus compuestos químicos, sus texturas, sus olores o sus sabores).

En el libro *The Flavor Bible* («La Biblia de los sabores») tenemos una buena recopilación de afinidades entre ingredientes. Por ejemplo encontramos que el atún

graso, el jamón ibérico y el caviar crean una deliciosa combinación mantecosa y salada que se funde en la boca; o que la naranja, las aceitunas, el comino y la menta crean una mezcla fresca y jugosa. Este listado revela determinadas preferencias por combinaciones específicas de ingredientes, que pueden depender de la región geográfica o de la cultura. *The Flavor Bible* codifica unas 1.500 afinidades entre ingredientes únicas en numerosas culturas regionales. No obstante, sólo supone una mínima parte de todas las posibilidades existentes.

Hoy en día la gastronomía ocupa un lugar fundamental en nuestras vidas, y nuestras vidas se están digitalizando cada vez más. El 14 por ciento de la población de Estados Unidos se consideran *foodies* y los *millennials* están cambiando todo su panorama culinario. Actualmente el 50 por ciento de los *millennials* se consideran comensales críticos, a los que les gusta experimentar y crear sus propios platos, yendo más allá de las recetas tradicionales. Probablemente se inspiren en Pinterest y otras redes sociales, hagan buen uso de los mercados de productos frescos y de los productos de invernadero a los que recurren los científicos para sus experimentos, y usen algoritmos de sustitución de alimentos para crearse versiones veganas o macrobióticas de sus platos favoritos. A diferencia de sus padres o abuelos, son una generación más preocupada por la comida sana, por las propiedades de los ingredientes y de su procedencia. Es más, son amantes de las emociones, que buscan experiencias gastronómicas especiales, con sabores intensos y texturas extremas. Con las nuevas plataformas digitales y las redes sociales, los *millennials* están centrando su atención en la gastronomía y redefiniéndola.

En nuestro planeta conocemos un número finito de ingredientes comestibles. Los ingredientes siempre han sido considerados las unidades básicas de los alimentos.

Así pues, tenemos un modo finito de combinar ingredientes. No obstante, el número de posibles combinaciones de ingredientes es enorme, y puede considerarse prácticamente infinito desde la perspectiva humana. La cantidad de datos de que disponemos sobre nuestras preferencias alimentarias, el origen de los alimentos, sus propiedades (si son vegetales o animales, sus calorías, sus compuestos alérgicos, su valor proteínico), la preparación de los alimentos y el consumo de alimentos va en rápido aumento. Hoy en día disponemos de recetas en línea, bases de datos con los compuestos químicos de los alimentos, recetas de alta cocina y de cocina tradicional, críticas de restaurantes, valores calóricos, nuevas taxonomías y clasificaciones de alimentos, y muchas cosas más.

Combinando una gran cantidad de datos, el «alma de los datos» de la cocina, podríamos responder a un montón de preguntas: ¿Cómo está cambiando el consumo de alimentos en el siglo XXI y cómo lo están modelando los *millennials*? ¿Cómo puede ayudarnos esta nueva abundancia de registros a estudiar el modo que tenemos

de comer? ¿Por qué tienen éxito algunos platos? ¿Qué nuevos platos podrían crearse? ¿Cómo influye la disponibilidad de estos datos en nuestras experiencias culinarias, nuestros hábitos y preferencias? ¿Cómo podemos modificar las recetas para que se adapten a todas las restricciones dietéticas de una persona específica, pero sin alejarse del original? ¿Por qué algunos platos son considerados innovadores y creativos y otros no?

Usando los macrodatos y la IA podemos transformar ingredientes poco familiares en apetitosas recetas. La cocina convencional suele basarse en la intuición y en la puesta a prueba de las recetas. Normalmente, son la experiencia personal y las convenciones sociales las que dictan nuestras nociones de cómo combinar ingredientes. Con la inteligencia artificial podemos adoptar un enfoque diferente, usando la química y la física como punto de partida para el descubrimiento de combinaciones de ingredientes antes desconocidas, y adoptar algoritmos para ampliar las posibilidades de creación de recetas nuevas. Gracias a la ciencia y a los ordenadores, las combinaciones inesperadas de ingredientes pueden convertirse en la nueva norma, con combinaciones sorprendentes en recetas sanas y sostenibles.

Los *gourmands* de 222 países, obsesionados por el chocolate, amantes del café y locos por las gambas, comparten sus fotografías en redes sociales gastronómicas. Es una enorme cantidad, equivalente a 62.000 fotos al día. No es casualidad que las fotos compartidas por estos usuarios aparezcan cada vez más en el entorno científico como parte de macrodatos: al igual que en otros campos de la vida diaria (asistencia sanitaria, política, tráfico viario, etc.), el análisis de estas instantáneas permite extrapolar información fundamental para innovar y mejorar nuestro modo de vida.

Es el caso de los investigadores del Qatar Computing Research Institute, que pasaron meses recopilando datos relacionados con fotografías de comida colgadas en Instagram para identificar información interesante sobre las tendencias más placenteras del mundo: las de comer bien, por supuesto, por la gratificación gustativa, visual y olfativa que da al mismo tiempo.

Los estudios realizados a partir de estos análisis han aprovechado la oportunidad de examinar nuestros alimentos favoritos, así como las dietas por la que optan diferentes países y diferentes clases sociales.

En Argentina, por ejemplo, lo que más le gusta a la gente, además de la merienda de media tarde, es el dulce de leche, una crema hecha con leche y azúcar caramelizado. En Estados Unidos, muchos platos nacionales se han convertido en los alimentos preferidos por muchos: desde las gambas de la costa al omnipresente beicon, aunque también los tacos mexicanos y el sushi japonés. En Canadá les encanta compartir fotografías de poutine (patatas fritas, pero servidas con salsa y queso). Eso nos lleva a los grandes clásicos internacionales: la #pizza es el alimento en el que más gente piensa cuando quiere darse un capricho; el #café sigue siendo la bebida más colgada en las redes de fotos; y el #vino es el número uno entre las bebidas alcohólicas.

En los países del norte de Europa y Holanda, en particular, les gusta comer sano. Por otra parte, hay tres países que destacan por su costumbre generalizada de postear y etiquetar fotos de alimentos que tienen muy poco que ver con el control de las calorías, revelando su preferencia por las recetas más sabrosas: los países en cuestión son Brasil, Argentina y Francia, dominados, respectivamente por miles de etiquetas *#gordice* y *#gourmandise*. Al fin y al cabo, el experto en sabores más célebre de la historia, Jean Anthelme Brillat-Savarin, solía decir ya en 1700 que «el gusto por la buena comida es uno de los principales vínculos que une a la sociedad».

Los macrodatos y la IA también están cambiando la producción de alimentos y el modo de comer de la gente. Probablemente habrás oído decir que la IA va a cambiarte la vida. A primera vista, podría parecer una de esas exageraciones que hace la gente sobre los avances tecnológicos como forma de vender la idea, pero en el caso de los datos y la IA, estas afirmaciones tienen cierto peso significativo. En otras palabras, no es sólo un concepto teórico, sino algo que está tomando forma ante nuestros propios ojos. Pensemos en la comida. Algo tan simple como lo que comemos se está viendo afectado en gran medida por el uso de éstos y la IA, en muchos casos de un modo evidente. Los datos están cambiando lo que comemos, dónde comemos e incluso cómo comemos, y el impacto general que podría tener todo ello —no sólo en nuestra vida, sino también en la vida de todos los demás— es enorme.

Quizá uno de los efectos más notables de los datos sobre la comida, en general, se produce en las fases iniciales: durante el cultivo. Durante décadas, los granjeros se han visto obligados a adoptar nuevas tecnologías para poder mantenerse al día con el rápido crecimiento de la población. En consecuencia, muchos han empezado a aplicar soluciones basadas en los macrodatos en sus instalaciones. El objetivo final no es el de cambiar los tipos de cultivos, sino el de aumentar la producción. Se calcula que en 2050 vivirán en el planeta unos 9.000 millones de personas, por lo que la necesidad de ampliar las cosechas es urgente. Los macrodatos y la IA están contribuyendo a este fin de diversos modos. La tecnología de agregación de datos se ha impuesto rápidamente en gran cantidad de instalaciones agrícolas.

La agricultura de precisión es otro sector que está despertando mucho interés. Con el uso de la tecnología GPS y el análisis de datos, mide y responde la variabilidad del terreno y sus cultivos. El resultado es un uso más eficiente de recursos como los fertilizantes y el agua. Para ello se están usando también plataformas como TerrAvion. Al disponer de imágenes aéreas en tiempo real, los granjeros pueden mejorar sus prácticas de riego, lo que mejora las cosechas con menos gastos. Los resultados hasta ahora son impresionantes. En un estudio realizado en cultivos de soja se observó que los granjeros habían ahorrado hasta un 15 por ciento en recursos como fertilizantes y semillas, mientras que en otro se observó que

el uso de la agricultura de precisión reducía el uso de agua a la mitad, al tiempo que la producción aumentaba un 16 por ciento.

Pero eso sólo es el cultivo de los alimentos. Para llevarlos de la granja al plato también hay que utilizar un montón de datos. Con los recogidos por los sensores durante las fases de cultivo y transporte, la industria alimentaria ha conseguido reducir significativamente el volumen de alimentos que se echan a perder. También se recopilan datos de las estadísticas de transporte y de los niveles de descomposición, lo que indica el mejor momento para trasladar el alimento al mercado. Reduciendo la cantidad de alimento que se echa a perder, la industria alimentaria puede asegurarse que no se desperdicia tanto. Al mismo tiempo, los restaurantes usan esos datos para comprobar la calidad del alimento. Los análisis a tiempo real permiten comprobar el alimento antes de que se sirva para asegurarse de que mantiene un alto nivel de calidad y que no provocará enfermedades a los consumidores.

Aprovechando los datos y la IA a todos los niveles de la cadena de valor, las empresas alimentarias pueden beneficiarse del enorme potencial del medio digital para la creación de un valor sostenido. El uso de la tecnología digital puede ayudar a las empresas a aprovechar recursos de un modo más responsable para con el medio ambiente, y a tomar mejores decisiones sobre la procedencia de cada alimento.

El lenguaje de los sabores

La IA y los datos también están cambiando el tipo de comidas que podemos llegar a tomar. La exploración de nuevas combinaciones de ingredientes (afinidades) posibles en un plato puede depender de muchos factores —geográficos, de temporada, psicológicos, etc.—, pero lo importante es esto: ¿Estas limitaciones se deben a leyes específicas, algoritmos o patrones que condicionen nuestro modo de combinar ingredientes? Y preguntas como ésta han sido objeto de amplias investigaciones en los últimos años.

Por ejemplo, hay indicios de que, en algunas cocinas regionales, las combinaciones algorítmicas de ingredientes tienen cierta preferencia por los resultados ricos en calorías, como el nivel de grasas o de azúcares, que pueden satisfacer la necesidad humana básica de buscar fuentes de energía. No obstante, el valor calórico puede no ser el único impulso en la búsqueda de combinaciones de ingredientes «buenas». De hecho, hay otros factores que pueden jugar un papel importante en la búsqueda de nuevas combinaciones creativas de ingredientes, como el placer, la alegría, la sorpresa, la comodidad, etc. De momento no está claro si hay otros principios subyacentes a nuestra cultura alimentaria y qué principios dominan en cada momento. No obstante, la IA puede ayudarnos a responder algunas de esas preguntas.

El sabor es un «lenguaje» que cualquier amante de los placeres del paladar querrá dominar. Algunos chefs buscan ingredientes que tengan sabores limpios, penetrantes y acentuados. Otros quizá quieran construir un sabor rico y profundo para luego introducir un punto de acidez, especiado o herbal como contrapunto. Una vez dominas el lenguaje del sabor, puedes usarlo para comunicar, y convertirte así en un chef mejor. Algunas combinaciones extrañas de sabores han dado resultados sorprendentes, y hay toda una rama de los macrodatos y la IA que estudia los motivos de estos extraños pero perfectos maridajes y cómo encontrar otros parecidos (por ejemplo, fresas con cilantro; piña, queso azul y vino blanco; u ostras con kiwi).

Algunos científicos alimentarios —y también algunos chefs— sospechan que los ingredientes que tienen compuestos de sabor en común dan mejor resultado juntos que los que no. Un ingrediente perfecto servido solo puede ser algo extraordinario, se trate de una fruta perfectamente madura y dulce o de una sedosa loncha de *sashimi*. Pero en el mundo real los ingredientes perfectos son muy raros, y hay pocos alimentos cuyo sabor no se pueda mejorar con un pellizco de esto o de aquello. Un poco de azúcar hace que las fresas saquen todo su sabor. Un chorrito de lima destaca la dulzura del melón. Unas gotas de vinagre dan un sabroso contraste a las patatas fritas con sal. Comprender qué combinaciones potenciarán el sabor de lo que se cocina es uno de los conocimientos más importantes que puede llegar a dominar cualquier cocinero.

Científicos y chefs han conjeturado en los últimos años que un probable factor decisivo para conseguir buenas combinaciones de ingredientes podría ser lo que llamamos el principio o la hipótesis del maridaje de alimentos. La hipótesis del maridaje de alimentos, propuesta en un principio por François Benzi y Heston Blumenthal, postula que si dos ingredientes comparten importantes compuestos de sabor, hay muchas posibilidades de que combinados creen una combinación agradable. En los últimos años, esta hipótesis ha llamado la atención de aficionados a la gastronomía, chefs y científicos. Si el maridaje de alimentos es uno de los principios de base que explican nuestras preferencias de sabores, científicamente esto nos permitiría predecir y construir numerosas afinidades de ingredientes nuevas de éxito basadas en los sabores de los que se componen y en los nuevos algoritmos de la IA alimentaria.

Resulta bastante obvio que dos tipos diferentes de carne tendrán, probablemente, elementos comunes en su perfil químico de sabores, pero hay otros maridajes menos evidentes: el chocolate y el queso azul comparten al menos 73 compuestos de sabor, y algunos chefs atrevidos ya los combinan en pasteles y trufas. Lo mismo se puede decir del caviar y el chocolate blanco, que comparten altos niveles de aminos, un grupo de proteínas que se han descompuesto de su estado de aminoácido y que guardan ciertas similitudes con el amoníaco; o del café y el ajo, o de la mandarina y el tomillo, o del plátano y el perejil. ¿Reflejan nuestras recetas este concepto a un nivel más general? En otras palabras, ¿usamos con frecuencia pares de ingredientes

que tengan fuertes vínculos (es decir, que compartan numerosos compuestos de sabor) en esta red de sabores, o los evitamos?

Entre los aspectos principales de la esencia de cualquier ingrediente está su temporalidad, su sabor, su volumen, su función (salado, reconfortante, refrescante, etc.), su distribución geográfica, su peso y sus afinidades de sabor. Y eso puede crear un número de combinaciones enorme. El número total de combinaciones posibles de ingredientes es de al menos 10^{15} (mil billones, el número de estrellas de nuestra galaxia), con una media de ocho ingredientes por plato y un total de cientos de miles de ingredientes posibles. Eso, sin contar todas las posibilidades que surgen si empezamos a deconstruir esos ingredientes creando nuevas formas y estructuras (por ejemplo, una esfera gelatinosa verde que tenga el aspecto de una aceituna y que contenga aceite de oliva en lugar de un hueso). Incluso los niveles de concentración importan. Por ejemplo, dependiendo de su concentración, el compuesto bencilmercaptano puede recordar el ajo, el rábano americano, la menta o el café, y otro compuesto (el decanal) puede insinuar un sabor entre la naranja y el albaricoque. El mismo compuesto, en una concentración diferente, puede saber a curri o a jarabe de arce.

Con tal cantidad de potenciales combinaciones, probablemente un chef no tenga tiempo en toda su vida de cocinar ni siquiera una pequeña fracción de todas estas opciones creativas. Si un cocinero quiere encontrar una serie de ingredientes que presenten un buen equilibrio entre acidez, grasa, sal y dulzura, y ver qué sabores se combinan bien entre sí, la búsqueda puede ser eterna, pero la IA puede ser de ayuda, y recientemente los expertos en datos han encontrado curiosas conexiones entre los ingredientes que pueden facilitar el trabajo.

De hecho, no parece que el maridaje de alimentos nos dé la respuesta a la combinación perfecta. Un reciente estudio realizado por científicos de la Universidad de Harvard con IA y big data ha demostrado que este efecto de los sabores en común depende mucho de las culturas. Los platos estadounidenses combinan más ingredientes con sabores en común de lo que cabría esperar. Por ejemplo, en la receta de las gambas y cigalas con tomate, plato típico de Estados Unidos, hay dos ingredientes principales (las gambas y el tomate) que comparten el 1-penten-3-ol, mientras que la mozzarella, el parmesano y el tomate comparten el 4-ácido metilpentanoico.

Pero los platos del este de Asia y los del aclamado restaurante El Bulli (tal como hemos observado en nuestra propia investigación) no lo hacen. De hecho, parecen evitar específicamente estas combinaciones de sabores compartidos, y la tendencia es la contraria. Cuantos más sabores comparten dos ingredientes, más a menudo aparecen en la cocina estadounidense, y menos en las recetas de los chefs del este de Asia o en los platos de El Bulli. Y la misma propiedad se observa en los ingredientes claves que «definen» estas cocinas: ingredientes característicos de América del Norte como los lácteos, los huevos y el trigo tienden a compartir compuestos de sabor,

mientras que ingredientes característicos del lejano Oriente como la salsa de soja, el arroz, la cayena y el jengibre no los comparten. La cocina vietnamita es la que más sigue el principio de maridaje de alimentos y la japonesa es la que menos.

Estos resultados científicos son fascinantes, y desvelan parte de la magia oculta de la cocina. Y el maridaje de alimentos no es el final de la historia. Parece que hay mucho más. De hecho, con los macrodatos y la IA encontramos que el mapa de ingredientes que suelen presentarse juntos en un mismo plato tiene una serie de propiedades muy interesantes.

¡Resulta que el mapa de sabores de los alimentos es curvo, como el de internet! De hecho, algunas grandes combinaciones de ingredientes sólo se pueden explicar usando los mismos fenómenos que explican por qué algunos servidores de internet que proporcionan altas velocidades de conexión sólo pueden localizarse pensando en internet como un mapa curvo, no plano. ¿Eso cómo se explica? Podemos situar los ingredientes sobre un mapa y crear un mapa de macrodatos de alimentos, en el que los ingredientes estén próximos entre sí si comparten sabores similares. Si los ingredientes han aparecido juntos en muchos platos, se creará una conexión entre los ingredientes. Si dos ingredientes nunca se han usado en un mismo plato, no habrá conexión entre ellos. Pero si los ingredientes comparten numerosos sabores, la conexión entre ellos será grande y fuerte.

Sería como dibujar la constelación de los ingredientes y sus relaciones. Un gran mapa de carreteras y autopistas que conectara los ingredientes a través de sus sabores con grandes líneas que conectaran los ingredientes que comparten numerosos sabores o que aparecen en los mismos platos. Si estudiáramos ese mapa, construido a partir de todas las combinaciones de sabores usadas en todos los platos creados en la historia, resultaría que dos ingredientes situados lo suficientemente cerca en línea recta quizá no estarían tan buenos juntos.

No obstante, si usamos un mapa curvado, podemos encontrar nuevas combinaciones de ingredientes que parecen alejados, pero que de hecho están cerca el uno del otro y que juntos pueden crear grandes combinaciones. Eso significa que a veces dos ingredientes que parece que no encajan, o que distan mucho el uno del otro, de hecho podrían combinar bien si se incluye algún ingrediente intermedio que los «una», que reduzca la distancia que los separa en ese mapa curvo imaginario.

Por ejemplo, ese mapa podría decirte que el pollo y el queso feta, que están alejados el uno del otro en el mapa de ingredientes, combinan bien si se les añade aceite de naranja, menta y jengibre. O que el pan y el jamón pueden combinar mejor con un poco de bacalao, romero e hinojo. Ésa es la belleza de la IA y de los algoritmos de datos, que pueden revelar la magia oculta tras muchas cosas.

La IA y la gastronomía

Mientras hacía mis estudios de doctorado, mi tutor se trasladó de Estados Unidos a Francia para colaborar en un nuevo laboratorio creado por el Instituto Federal Suizo de Tecnología, y yo me trasladé con él. Me mudé a Antibes, en el sur de Francia, el lugar donde Picasso había establecido su taller de escultura tiempo atrás.

Había aterrizado en la bella Provenza, con sus encantadores pueblecitos, sus espléndidos campos de lavanda y girasoles, sus pintorescas bodegas, sus jabones hechos a mano y sus aceites de oliva de prensado en frío llegados directamente del molino, sus dinámicos mercados y una pasión por la comida que yo no había experimentado nunca. Estaba en medio de lo que hasta entonces era el templo gastronómico de la cocina del mundo, Francia, un lugar donde aprendí a apreciar la cultura de los cultivos, de los platos de temporada, de cocinar con ingredientes simples y de hacer platos que llevaran a la mesa el sabor del sol y del Mediterráneo. Un lugar que —poco podía imaginármelo— marcaría mi evolución futura.

La *cuisine niçoise* es la cocina de la Riviera Francesa, una cocina mediterránea deliciosa, y esa zona, en términos culinarios, es una pequeña Meca de la gastronomía. Está llena de enclaves con sus propios platos —que no se hacen en ningún otro sitio— y de caóticas *vielles villes*, llenas de callejuelas sinuosas y tiendecitas. Se pueden encontrar muchas cosas que no se encuentran en otros sitios: *soccas*, *beignets*, *petits farcis*... Las *soccas* son crepes de harina de garbanzo gigantes, pero se hacen con aceite de oliva y son saladas, crujientes por los bordes y tiernas por el interior. Aprendí sobre sus tradiciones en la cosecha y sobre los ritmos de la *cuisine du soleil* que, al igual que las cocinas española, griega o italiana, está cargada de pasión y de salud. También aprendí a cocinar con flores, por ejemplo usando un cesto de capullos de rosa oscuros, que se pueden encontrar en puestos de mercado junto a especias del norte de África. Creo que me enamoré de la comida al momento. ¡Qué maravilla!

En el año 2000 acabé mi doctorado y me trasladé a California para seguir colaborando con mis colegas del sector de internet. Me instalé en Palo Alto, en pleno Silicon Valle, donde el Padre Palou, uno de los frailes franciscanos españoles que había fundado la Baja California, había plantado una cruz para indicar el lugar donde se construiría una misión bajo una secuoya gigante, el Palo Alto. Hoy en día Palo Alto es el corazón de las empresas emergentes de Silicon Valley, el lugar de nacimiento de Google o Facebook y sede de la famosa Universidad de Stanford, creada según se dice por Leland Stanford, magnate, político y buscador de oro, en las 260 hectáreas de su Rancho San Francisquito.

En California yo trabajaba en Inktomi, *startup* surgida de la Universidad de Berkeley que proporcionaba el motor de búsqueda por defecto de la mayoría de sitios web en aquel momento y una infraestructura que permitía redimensionar la web a gran velocidad. Era el mayor motor de búsqueda del mundo hasta la llegada de Google, y ofrecía servicios a Microsoft, AOL y Yahoo! Yo dirigía un equipo encargado de llevar internet al móvil, de modo que la gente pudiera navegar desde su teléfono. Y un día, mientras creaba nuevos algoritmos, analizaba montones de datos

de uno de los primeros nódulos celulares de internet del mundo, en el laboratorio de Ericsson en Berkeley, recibí un email de un joven periodista que me preguntaba si podría entrevistarme para un periódico local dirigido por la comunidad tecnológica latina de Silicon Valley.

Respondí afirmativamente, y nos encontramos en el University Café, sito en el 271 de University Avenue, en Palo Alto. La entrevista fue una agradable conversación; hablamos de las decisiones que había tomado hasta llegar a California, cosas como qué quería ser cuando fuera mayor a los diez años, cuál había sido mi primer trabajo, qué consejos daría para no quedarse atrás, o qué era lo que me había ido bien a mí hasta el momento. En un momento dado pasamos a una sección que él llamó «Confesiones profundas», y me planteó una pregunta bastante curiosa.

Quería saber quién era la persona que más me gustaría conocer. De hecho era algo en lo que yo nunca había pensado. Tuve que pararme a pensar. Podría haber dicho Bill Gates, Steve Jobs o Vint Cerf, el padre de internet. Pero en lugar de eso, tras una breve reflexión, respondí: «La persona que más me gustaría conocer es Ferran Adrià, el chef de El Bulli, el nuevo Picasso de la cocina. Me encantaría ser testigo de su fabulosa creatividad. Quizá pudiera sacarle un pellizco de inspiración para mi investigación científica...».

A lo largo de los años he aprendido que tienes que tener cuidado con lo que deseas... ¡porque puede que ocurra! Probablemente El Bulli haya sido el restaurante más famoso de todo el planeta. Tenía una temporada limitada: la de 2010, por ejemplo, iba del 15 de junio al 20 de diciembre. El día después de cerrarse la temporada ya estaba reservado para toda la temporada siguiente. Sólo podía acoger a 8.000 comensales en una temporada, pero recibía más de dos millones de solicitudes. El precio medio de una comida era de 250 euros. La revista *Restaurant Magazine* nombró El Bulli número uno de su lista de los 50 mejores restaurantes del mundo nada menos que cinco veces, en 2002, 2006, 2007, 2008 y 2009, y fue el número 2 en 2010.

En 2008, uno de mis mejores amigos científicos —un investigador indio que vivía en Nueva York— estaba trabajando a fondo para conseguir que la web funcionara mejor, escribiendo nuevos algoritmos, protocolos y diseños, e innovando en el uso de *big data*, analizando grandes series de datos para entender el rendimiento de la web y cómo mejorarlo. Pero también era un gran amante de la gastronomía y hacía tiempo que estaba obsesionado con comer en El Bulli.

Para conseguir una mesa, se apuntó al Catalan Institute of America en Nueva York, organización dedicada a reforzar los lazos entre Cataluña y Estados Unidos, empezó a tomar clases de catalán y a aprendió a bailar la sardana, el baile típico catalán. Esperaba conocer a algún lugareño que le ayudara a conseguir mesa en el Bulli... ¡y lo consiguió! Me enteré a través de un email que decía «*El Bulli i la mare que el va parir*» («El Bulli y la madre que lo parió»), y pensé: «¿Un ingeniero

informático que hace tantos esfuerzos para probar la cocina de Ferran Adrià?». Ahora sí que estaba realmente intrigado.

Unos años más tarde, el 30 de julio de 2011, el día que el aclamado restaurante El Bulli anunció que iba a cerrar, la noticia llegó a la portada del *Financial Times*. Y ese mismo día yo estaba escribiendo un artículo para una revista española, explicando que un equipo de científicos de datos e ingenieros de Telefónica R&D iban a colaborar con Ferran Adrià y El Bulli para combinar tecnología y cocina.

¡Así era! Tras más de quince años viviendo en el extranjero, había vuelto a Barcelona. La segunda cosa de la que más se hablaba después del FC Barcelona y Guardiola era El Bulli. Y cuando no hacía ni dos semanas del cierre del restaurante, Ferran Adrià declaró su intención de crear la El Bulli Foundation, un centro de innovación aliado con la tecnología digital que redefiniría la alta cocina para poder ofrecer a las nuevas iniciativas creativas una guía para la innovación.

La relación de Adrià con Telefónica empezó en noviembre de 2010, cuando fue contratado como embajador de la compañía. Telefónica I+D, la sección de investigación y desarrollo de la compañía, le ofreció la posibilidad de explorar cómo aplicar las nuevas tecnologías al mundo culinario para potenciar la innovación. Durante nuestro primer encuentro, mi padre, que había llegado a dominar el arte de las galletas de almendra, me envió un paquete para que pudiéramos probarlos en la pausa del café, y otro empleado trajo un pan de *sourdough* que había hecho él mismo para que lo probáramos.

Aquella mañana mezclamos comida y tecnología en una reunión muy animada, y enseguida me quedó claro que, del mismo modo que la última generación de chefs había innovado en el mundo de la gastronomía usando la química y la física, la nueva generación de chefs posiblemente integraran la IA y los datos en su evolución y en sus creaciones.

Cuando conocí a Ferran Adrià, le dije que tenía «alma de científico», tras observar su insistencia en publicar sus investigaciones y su metodología, que recordaba el proceso académico de revisión por pares que nosotros seguimos. El mundo de la ciencia y el de la gastronomía quizá usen procesos y metodologías similares, pero muy raramente se cruzan. En octubre de 2011 organizamos la primera conferencia de gastronomía y tecnología, en la que diversos ponentes hablaban de temas como la piratería gastronómica, el genoma de los alimentos, las recetas de código abierto, la gastronomía computacional o el desarrollo de lenguajes de programación gastronómicos.

Empezamos a adaptar sensores al equipo de los cocineros para comprender cómo se usaba y recorrimos el edificio. Para explorar interfaces gestuales, desarrollamos nuevos dispositivos conectados a una Kinect con los que los chefs podían consultar una base de datos en línea mientras estaban en plena elaboración de un plato, con las

manos sucias. También exploramos cómo usar sensores para seguir a las personas presentes en el edificio, para poder determinar la procedencia de un plato y ver dónde interactuaban los colaboradores: los usuarios podrían ver vídeos que mostraban la evolución de un plato y analizar el proceso y sus tiempos, para ver si se podía replicar.

Durante un tiempo tuvimos una impresora 3D en un rincón del laboratorio para conectar el reino digital con el mundo físico: los usuarios de la Bullipedia (la enciclopedia culinaria digital de El Bulli) podían descargarse archivos CAD e imprimirse sus propias maquetas gastronómicas. Produjimos moldes de chocolate —una cáscara marrón del tamaño de una nuez que se rellenaba de chocolate—, alimentos impresos en 3D, pero la aplicación más interesante de todas fue el diseño de herramientas que ayudaran a otros chefs y amantes de la gastronomía a encontrar su inspiración. Quizá el gran logro fuera llegar a comprender el poder de los datos y la IA como herramienta para encontrar nuevas recetas y ayudar a otros a crear las suyas.

Para ir más allá es necesario comprender lo que existía y lo que se había hecho hasta entonces en El Bulli, y que la Bullipedia fue creada por Ferran para crear una completa base de datos y conocimientos con una nueva clasificación de los platos, los ingredientes, las texturas, los olores y las elaboraciones. Si integramos los algoritmos y los datos a la comida, metáforas y conexiones que los chefs habrían tardado años en descubrir quedan inmediatamente al alcance de todos.

Hace unos años, los chefs habían empezado a acercarse a la física y la química para entender mejor la ciencia de la gastronomía. De hecho, en 1988 nació la «gastronomía molecular», como disciplina científica que estudiaba lo que sucede cuando cocinamos, y que ha ayudado a crear nuevos platos y nuevas cocinas. Por ejemplo, un huevo cocido exactamente a 68 grados centígrados tiene la textura de la crema, y un sorbete hecho en nitrógeno líquido tiene una espléndida textura aterciopelada; por otra parte, el procesamiento a alta presión (alternativa al calor como medio de cocción) permite extraer toda la carne de las langostas. Eso hace posible nuevas formas de cocina, como la deconstrucción de los alimentos, que consiste en descomponer cada ingrediente de un plato y transformarlo radicalmente, aunque preservando su esencia. En los últimos veinticinco años esta mezcla de disciplinas —gastronomía, ciencia, diseño, arquitectura y nuevos métodos creativos— ha llevado a la cocina a un nivel superior. Pero ¿qué podrían aportar los macrodatos y la IA?

En 2014 recibí una invitación para asistir a una reunión de la British Royal Society en Buckinghamshire. La Royal Society es una asociación que reúne a los científicos más eminentes del mundo, y la academia científica más antigua del mundo. Estaban creando un nuevo taller llamado «Gastronomía computacional: la creatividad y la comida en la Era de los Datos», que pretendía unir a los investigadores y a los profesionales más punteros de los campos de la ciencia

gastronómica, la ingeniería informática, la neurología, la sociología y la gastronomía para discutir sobre estas cuestiones e intercambiar ideas.

Los estudios más recientes realizados en este campo ya demuestran que la exploración de datos, su análisis masivo y la inteligencia artificial pueden darnos nuevas perspectivas sobre la elección de los alimentos, sugerirnos nuevas combinaciones de ingredientes e identificar factores culturales, antropológicos, psicológicos, gustativos, nutricionales y químicos que determinan qué alimentos nos gustan y por qué.

El taller tenía lugar en Chicheley Hall, una antigua finca de estilo barroco construida en el primer cuarto del siglo XVIII en la campiña inglesa. Entre los invitados estaban Harold McGee, científico gastronómico y escritor estadounidense autor de *La cocina y los alimentos: enciclopedia de la ciencia y la cultura de la comida*, obra que se ha convertido en uno de los libros más determinantes para toda una generación de chefs con mentalidad científica; Jonathan Lake, jefe de cocina del restaurante The Fat Duck del Reino Unido, conocido en todo el mundo por su interpretación científica de la cocina y por la integración de la percepción multisensorial en la experiencia gastronómica; y Michael Bom Frøst, director del Nordic Food Lab, fundado por el famoso restaurante Noma, que alcanzó el número uno en el mundo el año después del cierre de El Bulli.

Anteriormente ya se habían organizado talleres similares para impulsar el movimiento de la gastronomía molecular y física. Por ejemplo, el taller celebrado en 1992 en Erice (Italia) había reunido a científicos gastronómicos, químicos y cocineros profesionales para discutir sobre la ciencia y la física en que se basan las recetas tradicionales. Pero en este taller celebrado en Buckinghamshire en 2014 el grupo de asistentes y sus especializaciones eran bastante diferentes, aunque algunos de ellos, como Harold McGee, ya habían estado presentes en la reunión de Erice de 1992.

¿Por qué habían invitado a ingenieros informáticos como yo? Muy pronto me daría cuenta de que la comida también se puede analizar mediante conjuntos de datos masivos y con nuevos algoritmos de inteligencia artificial. Con la ciencia de los *big data* y la IA podíamos intentar entender por qué tenían buen sabor algunas recetas y otras no, por qué algunos ingredientes se combinaban entre sí con más frecuencia que otros, cuáles eran los patrones más populares de combinación de ingredientes por país, según cada chef o cada estilo de cocina, qué es lo que hace que un estilo de cocina sea único, crear nuevos platos, etc.

Los restaurantes pueden aprender mucho de los *datos y la IA*. ¿Qué pasaría si los dueños de un restaurante supieran que estaban sentados sobre una enorme mina de oro y que disponen de herramientas para explotar esa mina de oro? Esa mina de oro es el alma de los datos de un restaurante, que podría revelar patrones, tendencias y los vínculos con las operaciones necesarias para mejorar los resultados, la gestión del trabajo, las relaciones con los clientes, la reducción de los tiempos de espera y el

aumento de los márgenes de beneficio. Entre los recursos que podrían integrarse en el sector de la restauración hay una *app* de gestión que permite visualizar desde el teléfono datos en directo; una *app* de entrega de comida a domicilio para llegar a una clientela mayor, o una carta mejor y más atractiva.

Los macrodatos y la IA también podrían ser una herramienta útil para que el dueño de un restaurante sepa cuáles son los sabores más populares de su carta. Eso es importante, ya que a la larga querrá revisar la carta y eliminar los platos más impopulares, y añadir otros que tengan más éxito. Por ejemplo, Food Network y Wired idearon un estudio con el que demostraron que muchos platos recibían una puntuación mucho más alta cuando se les añadía beicon. Los análisis de *big data* e IA no sólo nos permiten saber el tipo de alimento que conviene introducir en la carta, sino que también permite valorar las iniciativas de promoción y el ajuste de precios consiguientes. ¡Cuanto más platos tengas en la carta que gusten a la gente, más posibilidades tendrás de contar con clientes felices que vuelvan una y otra vez!

Los macrodatos y la IA también podrían aprender de las recetas y combinaciones existentes, de lo que ha funcionado en el pasado y lo que no, y proporcionar a los chefs un conocimiento detallado de todos los productos e ingredientes (origen, composición, propiedades) para sacar lo mejor de cada ingrediente, destacar sus propiedades y al mismo tiempo crear y encontrar nuevas composiciones.

En Chicheley Hall, la casa de campo inglesa que albergaba el taller de gastronomía computacional, me invitaron a discutir alguno de los resultados que habíamos obtenido aplicando la ciencia de datos, que demostraba que el 35 por ciento de las combinaciones creadas por El Bulli son únicas y no se encuentran en ningún otro lugar del mundo, o que el número medio de aromas de los platos de El Bulli está justo por debajo del umbral impuesto a nuestra nariz por la naturaleza (hacia veinte olores simultáneos diferentes), a diferencia de otros platos de cocina tradicional, que a menudo incluyen más de cien aromas diferentes que nuestra nariz no puede apreciar al mismo tiempo (a pesar de que los humanos podemos discriminar más de un billón de estímulos olfativos). ¡Fascinante!

Pero los resultados más sorprendentes demostraban que aplicando la ciencia de los datos al mundo de los alimentos, y combinándola con lo que hemos aprendido de internet, usando mapas curvos a la hora de pensar en la gastronomía, los ordenadores podrían revelar nuevos patrones y crear composiciones nunca exploradas anteriormente. Dado el número de ingredientes que comemos los seres humanos, el número total de combinaciones posibles es muy grande, del orden de 10 elevado a la 15ª potencia. Y sin embargo, el número de recetas que comemos es de un millón más o menos, una fracción mínima de esa cifra. Lo que hace pensar que en nuestras recetas rige un principio de organización que, por decirlo así, separa el grano de la paja.

Así pues, los científicos gastronómicos siguen enfrentándose al desafío de descubrir las leyes que gobiernan las combinaciones de sabor y usarlas para crear

nuevas recetas para las papilas gustativas humanas. Estudiando los alimentos de diferentes culturas, Tiago Simas, científico de inteligencia artificial en Telefónica, ha descubierto que si analizas la red de sabores, el mapa que determina las combinaciones a las que se prestan para producir recetas interesantes, parece que las mismas recetas muestran un interesante algoritmo de combinaciones. La idea de base es que cuando dos ingredientes no tienen sabores en común es necesario un tercer ingrediente que tenga sabores en común con los dos.

Por ejemplo, el albaricoque y el whisky no comparten sabores entre sí, pero sí tienen sabores en común con el tomate. Eso crea una cadena de sabores que une los tres ingredientes, posibilitando su uso en la misma receta. El equipo de investigación de Tiago lo llama «puenteo de alimentos» (*food bridging*), y tiene un gran impacto en la recetas. Mientras que el maridaje de alimentos intensifica el sabor mediante la combinación directa en una receta de ingredientes con unos compuestos químicos similares, el *food bridging* suaviza cualquier contraste entre ingredientes. En América latina, por ejemplo, las recetas explotan tanto el maridaje de alimentos como el *food bridging*, mientras que las del lejano Oriente parecen evitar ambos principios. Las cocinas del sureste asiático, como la tailandesa y la vietnamita, parecen basarse sólo en el *food bridging*, mientras que la estadounidense y la de Europa occidental sólo usan el maridaje.

Desde luego el *food bridging* puede ser de ayuda para crear nuevas recetas. Pero una herramienta realmente universal para la creación de recetas tendrá que tener una amplitud mucho mayor para poder incorporar todos estos factores. Eso requerirá un trabajo significativo. No obstante, los científicos gastronómicos están aprendiendo cómo hacen los humanos para filtrar todas las combinaciones posibles de alimentos que acaban produciendo las que acabamos comiéndonos.

¿Y si el próximo gran chef fuera un supercomputador que usara datos masivos y la IA? ¿Y si las herramientas que estamos desarrollando con El Bulli pudieran condicionar el modo de cocinar de los jóvenes chefs y los gustos de los nuevos gourmets? La nueva generación de chefs parece estar deseando probarlo. De hecho, ante mis propios ojos, en la campaña inglesa de Cambridgeshire, Jonathan Lake, jefe de cocina del Fat Duck nos hizo una demostración de aplicación de los algoritmos científicos de *big data* y IA, con los que fue capaz de combinar melocotón con whisky (que no combinan bien individualmente y que de hecho se encuentran muy alejados el uno del otro en ese imaginario mapa estelar curvado de ingredientes), aunque un trozo de tomate seco hizo de puente y los conectó, creando una nueva combinación divina, elaborada no en una cocina sino en un ordenador.

Hace poco le pregunté a Ferran Adrià cómo pensaba él que podría ayudar la ciencia de los macrodatos y la IA a la creación de nuevas recetas. Adrià se giró hacia los que allí estábamos y dijo: «Nombradme un ingrediente al azar; uno cada uno». «Merluza», «tomate» y «vino», dijimos nosotros tres. Ferran nos miró y dijo: «¿Qué tal un ravioli de tomate, merluza y vino?» Y así, ante nuestros propios ojos, acababa

de crear un nuevo plato con ingredientes al azar, demostrando que la creatividad del genio no puede quedar eclipsada por la inteligencia de un ordenador.

Hace décadas que sabemos que hay problemas que los ordenadores nunca podrán resolver. Problemas para los que no se puede escribir un programa y que la tecnología *big data* y *IA* no puede resolver. Son problemas irresolubles. Hace tiempo que sabemos que los ordenadores tienen sus límites. Que hasta ahora las computadoras han afrontado problemas «sencillos», fáciles de computar. Pero el mundo es mucho más complicado de lo que pensábamos. Hay un mundo más próximo al de la creatividad y la imaginación donde las preguntas son más importantes que las respuestas y donde no todo se puede computar.

Así pues, sí, los ordenadores distan mucho de la perfección, y los datos masivos y la *IA* no tienen soluciones para todo, pero puede que un día, quizá menos lejano de lo que creemos, los robots y los ordenadores aprendan a crear burritos mejores de los que hacen los humanos. Y cuando llegue ese momento, la tarde en que nos estemos comiendo un burrito algorítmico, nos sentiremos a añorar los días del apogeo de la intuición humana, soñando con los aromas de la Provenza, o la sensibilidad y la inteligencia de la cocina de Ferran Adrià. Pero esta vez será el nacimiento de la inteligencia e intuición de las máquinas la que nos llevará en ese viaje de disfrute y placer, liderada por los datos y la *IA*.

Apostarlo todo a la ciencia de datos y a la IA

Confiamos en Dios. Todos los demás deben aportar datos.

W. EDWARDS DEMING

Un paseo por La Boqueria

Hace un día radiante en Barcelona y estoy a punto de perderme por las calles de esta bonita ciudad. Antes de conocerla tenía una impresión vaga de Barcelona como destino turístico y como sede de los Juegos Olímpicos a principios de los años noventa, cuando Freddy Mercury estaba en su momento álgido, y ahora le oigo cantando a todo pulmón la letra de aquella canción que describía Barcelona como un bello horizonte.

Estoy de acuerdo con él; la ciudad es un bello horizonte y un lugar vibrante que sin duda fue objeto de considerables inversiones tras las Olimpiadas, pero que aun así conservó la vida de sus calles, algo ruidosas pero a la vez serenas. El contraste es muy visible en las esquinas de antiguos edificios regios con bares de tapas para comer en el ambiente más desenfadado que he visto nunca. Viendo el paisaje urbano y la ubicación de Barcelona, ciudad del sur del Mediterráneo con un clima suave y un bonito litoral en el que se amarran junto los superyates y los barquitos de vela más pequeños, es fácil ver por qué tanta gente de todo el mundo acude a pasar aquí sus vacaciones de verano.

La cultura de la ciudad muestra una increíble ambición, un espíritu pionero que se remonta a los viajes de Colón a América, por no hablar de Pablo Picasso, descrito como uno de los artistas más influyentes del siglo xx. La influencia de Barcelona se mantiene en la actualidad, con una alcaldesa «de las más radicales del mundo», que es como han descrito a Ada Colau.

Sin duda la ciudad perdió algo de impulso durante la crisis financiera de 2008 y está atravesando por un importante proceso identitario, y aún pueden encontrarse rincones en la ciudad que parecen estar por acabar. Sería justo decir que es más bien una obra en curso, pero se podría decir también que, al igual que la Sagrada Família, es atrevida, bella, vanguardista y seguirá creciendo sin perder su elegancia.

Pero hay algo más; está en el ADN de la ciudad, en el corazón de su cultura. La historia de la exploración de nuevas fronteras, del arte y de los aspectos más humanos

de las cosas. Son los contrastes entre las tradiciones y el nuevo pensamiento que he descrito antes, el jaleo que se combina con un ritmo de vida tranquilo. Es la gente, la diversidad.

A mí me encanta comprar en el mercado de La Boqueria. Es una de las grandes atracciones turísticas de Barcelona, con su entrada en La Rambla, muy cerca del Liceu, el teatro de la ópera de Barcelona. Probablemente sea una de las mejores experiencias auditivas, olfativas y de color de toda Europa, llena de puestos de ricas frutas y verduras variadas, una variedad aparentemente ilimitada de pescados y mariscos, carne, quesos, embutidos y dulces.

Me gusta especialmente buscar erizos de mar de las agitadas aguas del Cantábrico, que baña la costa norte de España, donde nací yo, y guisantes *garrofal* y *floreta*, que son dos de las variedades más apreciadas de guisantes de la comarca del Maresme.

Según algunas crónicas, ha habido mercado en este lugar desde 1217, y aunque hoy en día es una atracción turística de pleno derecho, siempre ha sido el lugar donde iban a comprar los lugareños. Muchos de los grandes restauradores y chefs de Barcelona compran aquí sus ingredientes, lo que dice mucho de la calidad de los productos en oferta.

A pesar de todo, actualmente no resulta fácil atravesar la entrada principal, llena de grupos de turistas curiosos, para indicarle al pescadero el pescado que quieres o el tentador pedazo de queso de cabra asturiano que te vas a llevar, y en muchos casos hay más carteristas de lo que cabría esperar.

Yo vivo cerca, en la animada calle comercial de Portaferriça. La «porta ferriça» era una de las entradas de la segunda muralla de Barcelona, construida en el siglo XIII. En uno de sus extremos se encuentra una fuente que en principio era abrevadero para los caballos que circulaban por la zona, y por ello la mayoría de las casas aún tienen una entrada con patio para la entrada de coches de caballos.

Aquí es donde se encuentra el Palau Palmerola, la casa del conde de Fonollar, que vivió en este lugar durante el siglo XVII, y que actualmente contiene talleres de pintores, fotógrafos, diseñadores, chefs, y donde yo alquilo un piso. El edificio fue abandonado a su suerte durante largos períodos de tiempo, hasta que el famoso fotógrafo Outumuro y otros decidieron rescatarlo. Entre los hallazgos más sorprendentes que hicieron hay una serie de frescos de 1784 obra del pintor Pere Pau Montaña en lo que era el salón de baile del conde.

Cuando me mudé a este lugar el barrio era un lugar concurrido pero acogedor. Podías ir a pie al mercado de La Boqueria, hacer la compra y traerte los productos frescos a esta casa del siglo XVIII del Barrio Gótico. No obstante, en los últimos años las cosas han cambiado mucho. La crisis azotó España, y la ciudad ha ampliado muchísimo la oferta para turistas. Hasta hace poco se podían encontrar cientos de turistas paseando por los alrededores de La Boqueria, tomando fotos con sus teléfonos móviles y disfrutando de todo lo que ofrece la ciudad. Los frescos de

Montaña también se han convertido en destino de diversos circuitos turísticos, y han empezado a formarse colas en el exterior del edificio para poder mostrarlos a los visitantes, ya que la oferta turística se está volviendo cada vez más diversificada.

Que Barcelona es una de las grandes ciudades del mundo es algo en lo que la mayoría estamos de acuerdo. La arquitectura, las fiestas, la gastronomía... Es una ciudad dinámica y con alma. Pero el aumento del turismo había creado un problema entre los vecinos, que se quejan de las limitaciones de movimiento que el turismo crea en el mercado y en las calles adyacentes, del ruido y de que no pueden desarrollar sus actividades diarias como antes.

En septiembre de 2014 los vecinos de Barcelona, hartos de las hordas de turistas que llenaban sus calles, llevaron su protesta contra el turismo masificado a uno de los lugares más emblemáticos de la ciudad: la Sagrada Família. Unas trescientas personas organizaron una estentórea manifestación por la zona para quejarse de cómo les arruinaban la vida los visitantes, y un documental titulado *Bye Bye Barcelona* recogió las opiniones de los lugareños y sus quejas por el impacto negativo del turismo.

El número de turistas que visita Barcelona se ha triplicado o más en los últimos veinte años, y no se prevé que la progresión vaya a frenarse de forma drástica a largo plazo. Las cifras de 2013 demostraron que Barcelona superó los 7,5 millones de turistas anuales por primera vez en su historia, con franceses, británicos y estadounidenses a la cabeza del ranking de visitantes.

Sin embargo, hace un par de años, cuando entré en el mercado un sábado por la mañana, observé que había pasado algo. Había algo diferente: por unas horas, fue un mercado de La Boqueria completamente diferente, como si hubiéramos retrocedido en el tiempo, que ofrecía una escena que no había visto en años. En los puestos del mercado y en los pasillos había menos turistas y muchos más lugareños haciendo sus compras.

Cuando salí a La Rambla eché un vistazo a uno de los periódicos y leí: «Barcelona prohíbe la entrada de grandes grupos de turistas al famoso mercado de La Boqueria para que los lugareños puedan hacer sus compras». Los vecinos se quejaban de que la ciudad estaba invadida por los turistas, y el ayuntamiento quiso hacer algo al respecto, así que prohibió la entrada de grupos de más de quince turistas al mercado en las horas de mayor actividad —de las 8.00 a las 15.00 los viernes y sábados—, autorizando a los guardias de seguridad del mercado a acompañar a los grupos mayores a la salida del mercado en caso necesario. Los vendedores del mercado siguen quejándose de que hay que hacer más, pero al menos es un paso positivo.

No obstante, si eres alcalde de la ciudad, ¿cómo haces para decidir estas restricciones? ¿Cómo sabes qué tipo de turistas hay en cada lugar y en cada momento, qué vecinos son los más afectados y qué habría que hacer para que todo el mundo quede contento? ¿Cómo sabes que tienes que evitar la entrada de grupos de

más de quince turistas los viernes y sábados por la mañana para mejorar la experiencia de la compra de alimentos en los mercados locales? ¿Qué nacionalidades de turistas sufrirán más el impacto? La respuesta está en el alma de los datos de la ciudad, y en la IA.

La IA y el turismo

Las investigaciones del Consejo Mundial de Viajes y Turismo han demostrado que los datos masivos y la IA están transformando el sector turístico global. ¿Cuántos turistas del Reino Unido visitan Barcelona en ocasión de un partido de fútbol? ¿Cuál es la composición (por nacionalidad) de los turistas que visitan una región vinícola específica? ¿Cuánto dura de media una visita? ¿Ha sido efectiva la última campaña turística en el extranjero? ¿Qué conclusiones se pueden sacar sobre el turismo y el atractivo de dos regiones diferentes comparándolas? Hay que basarse en la información, no en cálculos aproximados, y para ello no hay más que usar datos acumulados y anonimizados.

Los datos de uso de las aplicaciones móviles son sensores de la sociedad y hoy son omnipresentes. Individualmente, estas señales anónimas no dicen mucho, pero de forma combinada pueden usarse para hacer un retrato de las calles. Bien procesada, esta información puede aportarnos datos continuados sobre diversos sectores: turismo, planificación urbana, gestión del transporte... Hoy en día, los consejos turísticos y las autoridades de planificación urbanística se basan en estudios manuales, entrevistas a nivel de calle con expertos realizadas por el Ministerio de Industria, Energía y Turismo, las principales organizaciones públicas y de seguridad, o por empresas privadas. El problema es que esos resultados proceden de cuestionarios hechos a mano que salen caros y que en el mejor de los casos se pueden aplicar varias veces al año, con poca granularidad en el tiempo y el espacio, lo que da unos datos de calidad bastante baja sobre el comportamiento de los turistas en la ciudad.

En la mayoría de casos, la información sobre los turistas se obtiene a partir del número de vehículos que cruza la frontera, pero eso deja de lado los que llegan por barco o en avión. En otros casos, las aerolíneas, las empresas de ferrocarril y las autoridades portuarias disponen de la información de los turistas que llegan a España por aire, por tierra o por mar. No obstante, aunque tienen datos exactos sobre los turistas que entran en el país, no la tienen sobre lo que hacen una vez llegan a destino.

Hace una década más o menos, los organismos reguladores hicieron obligatorio que todos los teléfonos móviles nuevos tuvieran implantada una tecnología de localización como el GPS para poder localizar a quienes realizaran llamadas de emergencia. Esa norma fue una piedra angular en la historia de los *big data*, porque

cambió la naturaleza de la información que se puede capturar y empezó a extraerla de las aplicaciones móviles, no ya de ordenadores fijos o portátiles.

El GPS y los sistemas de navegación hacen mucho más fácil encontrar el camino más rápido para llegar a los sitios, pero tienen un problema: envían a todo el mundo por la misma ruta, congestionándola. La gente que conoce bien una ciudad ya sabe la respuesta a esto: tomar una ruta alternativa. Los sistemas de navegación también conocen las rutas alternativas, pero no se plantean el problema desde un punto de vista global.

Eso es lo que plantea un estudio publicado en Nature Communications por Serdar Çolak y Marta González, del MIT, y Antonio Lima, de la Universidad de Birmingham, que registró datos de localización por GPS de miles de millones de aplicaciones móviles para analizar los modelos de tráfico en hora punta en Boston y la bahía de San Francisco (Estados Unidos), en Río de Janeiro (Brasil) y en Lisboa y Oporto (Portugal).

Lo que hallaron es que, sólo con que un puñado de conductores se dejara guiar por un modelo de enrutamiento «con conciencia social» —tomando una ruta ligeramente menos conveniente en lugar de sumarse a la mayoría y tomar las rutas más congestionadas (enrutamiento egoísta)—, podría reducirse el nivel total de congestión del tráfico hasta un 30 por ciento.

No obstante, los sistemas de localización por GPS gastan mucha batería, por lo que suelen estar apagados cuando la gente tiene el móvil en el bolsillo. Las aplicaciones móviles envían señales periódicamente a redes como WiFi mientras los usuarios se desplazan de un sitio a otro, sin necesidad de GPS, aunque no estén usando el teléfono, lo que proporciona un medio más eficiente para entender cómo se mueven los turistas por la ciudad, sin que eso suponga para ellos un gran consumo de batería. Analizando de forma anónima el número de aplicaciones activas por país mientras los turistas se mueven de un lugar a otro podemos aprovechar los datos de localización para sacar conclusiones sobre los principales países de origen de los turistas que nos visitan y que van a un destino determinado.

Para poder analizar los datos, tenemos que asegurarnos de que se protegen todos ellos sustituyendo información privada por un identificador asignado al azar que no permite identificar a ningún turista en particular. Además, éstos no se analizan individualmente, sino por grupos de personas (por ejemplo, por nacionalidad), de modo que queden anonimizados. Los datos de las aplicaciones móviles son como un censo o, por usar otra analogía, como unas elecciones. Puedes ver cuántos votos recibe cada candidato o cada partido, y cuál ha sido la «variación» desde las elecciones anteriores, pero nunca descubrir qué papeleta ha puesto en el sobre una persona en la intimidad de la cabina electoral. Se pueden extraer tendencias del conjunto de datos, pero no identificar a los individuos.

Poder monitorizar el flujo de personas que se mueve por una ciudad en tiempo real usando la IA proporcionaría una información valiosísima a urbanistas,

autoridades de transporte, ingenieros viarios e incluso a algunas empresas. Los horarios de los autobuses podrían tener en cuenta las variaciones horarias o diarias; los publicistas podrían saber qué carteles son más rentables. Esta información se puede recoger con los helicópteros de tráfico, con las cámaras exteriores, las patrullas de policía, con sensores integrados en las calles, con las unidades de rastreo de los vehículos, con los datos de los medios de transporte públicos y con las encuestas. Pero la imagen resultante en muchos casos es deficiente, cara o ambas cosas.

No es difícil pensar en otros usos para la IA y los datos de localización. Los agentes inmobiliarios, por ejemplo, podrían tener más fácil la valoración de las propiedades comerciales determinando cuántos peatones pasan por delante de un escaparate determinado. Los anunciantes querrán saber cuántos ojos pasan por delante de una valla publicitaria y cómo varían esas cifras entre laborables y festivos. Y las oficinas de promoción turística quizá quieran cambiar sus campañas en el extranjero cuando sepan qué nacionalidades (identificadas por sus operadores telefónicos) pasan más tiempo en la ciudad y cuáles prefieren echarse a tomar el sol en la playa.

Gracias a la IA y al alma de los datos de la ciudad, representada por los diferentes conjuntos de datos abiertos y los datos móviles anonimizados, hoy en día Barcelona es probablemente una de las pocas ciudades del mundo que dispone de mapas de movimiento de personas. Los mapas creados con los de localización por móvil darán a las autoridades de la ciudad un poder extraordinario y sin precedentes para cambiar los horarios de los autobuses de Barcelona según la variación de la demanda a lo largo del día. Eso también podría inducir mejoras en la colocación de la iluminación viaria, las señales y el trazado de las calles. Parece que los mapas de movimiento de personas tienen ante sí un gran futuro.

Las aplicaciones móviles son los nuevos sensores de nuestros tiempos, y dan una gran cantidad de información sobre la conducta humana. Cada día hay 800.000 personas en *roaming* fuera de su país. El turismo es un sector importantísimo en España, responsable del 5,7 por ciento del PIB, con una contribución total directa de 58.000 millones de euros en 2013, y el 91 por ciento de los móviles están a menos de un metro de los turistas las 24 horas del día. El turismo es uno de los sectores en el que más impacto han tenido. Y los datos de localización de las aplicaciones móviles pueden ser el nuevo canal para conectar el sector y los turistas, y facilitar que el sector turístico tome decisiones económicas y sociales más precisas y procesables.

Cuando trabajaba en las cercanías de San Francisco, compartía despacho con un ingeniero de software ruso, uno de los más brillantes que he conocido nunca y con quien aún conservo la amistad. En una de mis recientes visitas a San Francisco, fuimos a cenar a nuestro restaurante favorito en Silicon Valley, y le invité a que viniera a verme a Barcelona aprovechando que iba a dejar un trabajo con una *startup*

y empezar con otra. Me prometió hacerlo, y me preguntó si podríamos visitar Montserrat. «¿Por qué Montserrat precisamente?», me pregunté. No imaginaba que fuera uno de los lugares que mi amigo ruso pudiera querer visitar en su primer viaje a España, pero tampoco era la primera vez que me sorprendía.

De hecho, hace unos cuantos años sería impensable ver muchos turistas rusos en Montserrat. No obstante, cada año 150.000 rusos visitan el templo. Son la segunda nacionalidad en visitas a Montserrat, sólo superados por los españoles, y seguidos por franceses e ingleses.

Los rusos son católicos ortodoxos, y la basílica de Montserrat tiene un gran interés para ellos porque conserva una colección única de unos 170 iconos ortodoxos de origen griego, turco, rumano, serbio, sirio y ruso, expuestos con mucho gusto, que son una delicia para el visitante. La mayoría fueron comprados en mercados de segunda mano tras la disolución de la Unión Soviética.

Pero lo más fascinante es que los datos de los móviles y la IA ya daban esa misma información, una historia inesperada a la espera de que alguien la revelara. La historia de los macrodatos que indicaba que los rusos estaban a la cabeza del turismo religioso en España, que Sant Sadurní era la siguiente escala tras Montserrat, para descubrir la región del cava, y que normalmente de ahí se pasaba a las playas de Salou. Pero la historia de otros turistas era muy diferente.

Por ejemplo, los argentinos, los brasileños y los portugueses prefieren visitar Madrid, mientras que Barcelona recibe más visitas de los países del norte de Europa, ¡y sólo el 4 por ciento de los turistas visitan ambas! Los turistas permanecen de media 2,2 días en Barcelona, y visitan sobre todo los barrios de Ciutat Vella, el Eixample y Sant Martí, mientras que en Madrid se concentran en el centro. Los hoteles de cinco estrellas suelen atraer a los asiáticos, árabes, suizos y rusos. Los de cuatro estrellas son visitados por los ciudadanos del resto de Europa, salvo por los franceses, portugueses y suecos, que a menudo van a hoteles de tres estrellas. ¿No sería fantástico que los dueños de los hoteles dispusieran de toda esta información para que pudieran personalizar su oferta gastronómica, sus paquetes y otras necesidades?

La mayoría de turistas llegan a Barcelona en barco, con los grandes cruceros, y no por avión, como cabría esperar. Visitar la Sagrada Família se convierte en misión imposible de las 12.00 a las 15.00, y el Parc Güell se llena mucho de las 10.00 a las 12.00, lo que abre nuevas oportunidades a las guías turísticas personalizadas y a las *apps* turísticas basadas en la tecnología *big data* y la IA. Y si combinamos esos datos con los de gasto de las tarjetas de crédito, podemos ver que el gasto medio por noche es de 129 euros, y los que más gastan son los turistas de Singapur, China, Japón y los Emiratos, mientras que portugueses, franceses, daneses, suecos y argentinos son los que menos gastan.

Cada año, del 9 al 17 de mayo, Gerona celebra el «Temps de Flors», una de las fiestas florales más populares de Europa. Durante diez días, las calles de la ciudad se llenan de música y cobran vida con el color y los olores de las flores exóticas. Durante este período, miles de visitantes abarrotan la ciudad para disfrutar del encanto de esta fiesta típica catalana en una ciudad que se ha convertido en uno de los exteriores más famosos de la serie Juego de Tronos.

Durante dos años seguidos, el ayuntamiento de Gerona ha adoptado la tecnología de datos en este evento turístico para asegurarse el éxito de la fiesta.

Uno de estos estudios hizo posible que la ciudad de Gerona comprendiera el comportamiento de los turistas, además de registrar su procedencia y el tiempo de permanencia. Calculó, por ejemplo, que Gerona tuvo un total de 244.199 visitantes durante la fiesta, el 90 por ciento de ellos procedentes del resto de Cataluña, el 2 por ciento del resto de España y el 8 por ciento de otros países.

Recientemente, gracias a la IA y el análisis de datos, un grupo de científicos de Barcelona ha desarrollado un algoritmo de mapeo para encontrar el recorrido más pintoresco, usando los recogidos de las aplicaciones móviles. Nada que ver con la velocidad y con evitar el tráfico: ahora ya puedes seguir rutas que «no sólo son breves, sino también agradables emocionalmente». De hecho, los científicos crearon una base de datos con gráficas para que el algoritmo fuera capaz de seleccionar los recorridos más pintorescos. Los investigadores manifestaron que esos recorridos tienden a ser un 12 por ciento más largos que las rutas más cortas (la diferencia es de unos minutos), pero que resultan más agradables, y que crean mapas de calor de rincones bonitos de la ciudad, serenos y alegres.

La profesora Sandy Pentland, del MIT, también ha desarrollado un modelo científico de redes y un concepto de sociedad condicionada por los datos en la que se pueden usar datos muy baratos, como los de los teléfonos móviles, para monitorizar el estado de una ciudad, de forma continuada y en tiempo real. Por ejemplo, describe la capacidad de poner en relieve los mapas de pobreza, de mortalidad infantil, los índices de criminalidad, los cambios en PIB y otros indicadores sociales, actualizados a diario y barrio por barrio, usando la teoría de redes y los datos de los teléfonos móviles.

En el caso de Barcelona, los expertos en datos han demostrado que aproximadamente un 30 por ciento de los turistas se quedan en el crucero para disfrutar de sus instalaciones y que no salen del barco para ver la ciudad. Da la impresión de que hay mucho margen de mejora si se quiere atraer a la gente a la ciudad y hacerlo de un modo coordinado. Pero en el caso de los que bajan del barco y recorren la ciudad, cuando anochece vuelven a recorrer La Rambla en dirección al puerto y se recogen, dispuestos a partir en dirección a un nuevo destino.

Al irse se llevan sus teléfonos móviles, y con ellos el rastro que dejan en la ciudad. Pero no siempre es así: algunos teléfonos permanecen en la ciudad por la noche, incluso después de que el barco vuelva a partir, lo que indica los turistas distraídos que con tanta compra o tanta diversión han perdido el barco, y unos cuantos teléfonos móviles extranjeros permanecen activos en la ciudad incluso días o semanas después de la partida del barco, y con el tiempo acaban concentrándose en un barrio determinado de la ciudad: La Mina.

La IA y la seguridad

¿Es segura Barcelona? Sí. Barcelona es muy segura. En Barcelona se producen poquísimos crímenes violentos, y pasear por la ciudad, tomar el metro y el taxi no tiene prácticamente riesgo. De hecho, el «Índice de Ciudades Seguras de 2015» creado por la Unidad de Inteligencia de *The Economist* confirma que Barcelona es una de las ciudades más seguras del mundo. El informe de 2015 concluye que Barcelona es la tercera ciudad más segura de la Unión Europea, y la cuarta de Europa, sólo por detrás de Estocolmo, Ámsterdam y Zúrich. En el ranking mundial, Barcelona es la 15ª ciudad más segura según *The Economist*. De modo que sí, Barcelona es segura.

Aun así, también es cierto que hay muchos carteristas en los puntos de mayor afluencia turística (en particular en Las Ramblas), en las estaciones de metro del centro y en la estación de ferrocarril de Sants. No en el resto de la ciudad. Hasta el final de los años sesenta, La Mina era poco más que una extensión de campos de cultivo, con ganado y aldeas dispersas en la periferia de Barcelona. Los altos niveles de marginación social, con un alto índice de analfabetismo, dieron una terrible reputación al lugar, con titulares de periódico como «La Mina: barrio sin ley» o «La Mina, zona peligrosa».

Hoy en día, este desgraciado legado ha dejado a La Mina un índice de marginación social que es el más alto de toda el área metropolitana de Barcelona. Sufre de un trazado urbano que ha creado calles encerradas en una estructura como de fortaleza, aislada del mundo exterior. La densidad de población y de ocupación de las viviendas es muy alto; las casas son de muy poca calidad y ofrecen muy poco espacio vital. El barrio supera la media en cuanto a niveles de pobreza, y el índice de analfabetismo alcanza el 25 por ciento. El desempleo, el trabajo irregular y el absentismo escolar son muy altos. La comunidad ha sufrido una intensa degradación, con altos índices de criminalidad y una grave fractura social.

Y éste es el lugar donde van a parar un gran número de teléfonos móviles extranjeros mucho después de que se vayan los turistas, probablemente teléfonos robados a la espera de reciclarlos y venderlos en el mercado negro por una mínima parte de su precio, por ejemplo en Els Encants, mercadillo situado bajo el nudo viario

elevado próximo a la Torre Agbar (esa bonita estructura en forma de pepino azul y rojo hacia el norte de la ciudad).

La delincuencia crea patrones, al igual que todo lo que hacemos los humanos cuando lo observamos a una escala lo suficientemente grande. Así, aunque el comportamiento individual puede resultar difícil de prever, determinando el comportamiento medio de una población y luego relacionando a los individuos con ese patrón para determinar los «aciertos» podemos obtener resultados sorprendentemente precisos.

Un estudio reciente titulado «Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data» («Érase una vez un crimen: hacía la predicción de la delincuencia a partir de la demografía y los datos móviles»), de Andrey Bogomolov, Bruno Lepri, Jacopo Staiano, Nuria Oliver, Fabio Pianesi y Alex Pentland, analiza el uso de los datos de los teléfonos móviles, la inteligencia artificial así como los demográficos para predecir la delincuencia por zonas geográficas.

En el estudio usan registros de comportamiento agregados y anonimizados de una ciudad, derivados de la actividad de los datos móviles, para intentar hacer una predicción de la delincuencia. A diferencia de estudios anteriores, que han usado datos históricos sobre los perfiles de los delincuentes, los hallazgos de este grupo apoyan la hipótesis de que los datos agregados de comportamiento humano capturados desde las aplicaciones móviles, en combinación con una información demográfica básica, pueden usarse para predecir la delincuencia. En sus resultados experimentales con datos de delincuencia reales y de aplicaciones móviles en Londres, han obtenido una precisión de casi el 70 por ciento en la predicción de si una zona específica de la ciudad se convertirá en un barrio con actividad delictiva o no.

El estudio combinaba datos de Twitter de la zona de Londres, registros de éstos de aplicaciones móviles y una serie de conjuntos de datos de acceso público que incluía los casos delictivos, las ventas de viviendas, el transporte, la meteorología y los perfiles de los barrios de Londres en relación con el índice de personas sin techo, de residentes, el mercado inmobiliario, la situación económica del ayuntamiento y el bienestar social. Mediante la IA y empleando algoritmos de aprendizaje automático, los resultados demostraban que el número de personas que se quedaba en casa en una zona determinada era de gran importancia para determinar la criminalidad (cuanta más gente en casa, más vacías las calles y más posibilidades había de que se produjeran delitos).

También fue interesante observar el papel que jugaba la diversidad de personas en una zona para predecir los índices de criminalidad, la predictibilidad de una zona en cuanto a los patrones individuales de las personas que circulan por la zona en el transcurso del día. Un lugar visitado por gran variedad de individuos a diario presenta un riesgo mayor de ser escenario de delitos que un lugar caracterizado por un patrón regular de individuos y por un flujo periódico de movimientos a lo largo del día. La

diversidad de los patrones en relación con los diferentes grupos de edad, los diferentes usos del lugar (casa o trabajo) y en relación al género demostraba ser un buen factor de predicción del nivel de delincuencia de una zona determinada. Curiosamente, otros conjuntos de datos, como los de Twitter o la información de dominio público sobre el barrio no sirvieron de gran ayuda para la detección de delitos, debido a su escasa precisión temporal y espacial. Tal como muchos han señalado, este fascinante trabajo y sus conclusiones recuerda vagamente la película *Minority Report*, en que se predecían asesinatos antes de que se produjeran por medio de visiones del futuro.

Los datos también están ayudando a la policía de Nueva York a resolver los crímenes más rápidamente. Por su parte, en Chicago también están usando algoritmos de aquéllos para prevenir el aumento de la criminalidad. Desde un ordenador de la comisaría central de policía de Chicago, en el South Side, el subcomisario Jonathan Lewin puede extraer una lista de unos 1.400 nombres de personas con antecedentes delictivos. Junto a cada nombre hay una foto carné, sus datos demográficos, su historial de arrestos y su afiliación a bandas. También hay una puntuación.

La puntuación, determinada por un algoritmo basado en 11 variables ponderadas, mide la probabilidad de que una persona se vea involucrada en un tiroteo, sea como víctima o como agresor. La puntuación va del 1 al 500, y cuanto más alta sea, más probabilidad tendrá esa persona de formar parte de una actividad violenta.

Chicago está a la cabeza de una tendencia que se está extendiendo entre los departamentos de policía de Estados Unidos, que usan los macrodatos y los algoritmos para determinar las personas y los lugares con mayor riesgo de violencia. Algunas de estas previsiones de riesgo son sorprendentemente precisas. ¿Pueden salvar vidas la IA y las matemáticas? Hace tiempo que se usa la vigilancia policial predictiva para determinar los lugares con mayor riesgo de actividad criminal, pero ahora estamos pasando al nivel siguiente.

China, que en su día era el reino de las películas de ciencia ficción, está usando sus sofisticados sistemas de vigilancia y de reconocimiento facial como base para un nuevo «entorno de información unificada» que se usará para predecir si alguien va a cometer un crimen antes de que lo haga realmente. Por orden del gobierno, la agencia de defensa estatal más grande del país está construyendo una plataforma de software analítico por IA que podrá cruzar información de cuentas corrientes, trabajos, aficiones, patrones de consumo y filmaciones de las cámaras de vigilancia para identificar a potenciales terroristas.

Imagínate un mundo en que los agentes de policía, provistos de gafas de realidad aumentada, sean capaces de identificar a cada individuo de una multitud. Con la identificación de los individuos, todos sus datos, desde su historial médico a su actividad en las redes sociales, datos demográficos y antecedentes policiales afloran y

pasan por un algoritmo, que permitirá identificar al potencial autor de un delito, con lo que la policía podrá tomar medidas. En China, este panorama al estilo «Robocop» ya no es cosa de un futuro lejano.

Recientemente se ha hecho un gran estudio con IA y aprendizaje automático para determinar si sería posible usarlos como ayuda a los jueces en las vistas por violencia doméstica, por ejemplo para tomar la decisión de conceder la libertad provisional a un individuo o no. El estudio examinó 28.000 vistas por violencia doméstica durante el período 2007-2011. Luego se comprobó si esos 28.000 individuos habían reincidido o no en los dos años siguientes. La máquina examinó 35 características, como los cargos y condenas anteriores, o datos demográficos como la edad y el género, para determinar la posibilidad de reincidencia. Pese al número relativamente pequeño de campos de datos usados, los resultados fueron impresionantes. Los tribunales suelen conceder la libertad bajo fianza en la mayoría de esos casos tras la lectura de cargos, pero se demostró que el 20 por ciento de los acusados en libertad reinciden en los 24 meses siguientes. El análisis demostró que aplicando el modelo se podría haber refinado la selección, con un índice de reincidencia de sólo el 10 por ciento.

La IA y las finanzas

Las tecnologías de IA y *big data* están desafiando el modelo de banca tradicional, incluso más allá de las criptomonedas. Todo el mundo sabe que la banca es un sector competitivo, pero lo que está cambiando es la forma de esa competición. El campo de juego está cambiando gracias a la aparición de tecnologías disruptivas que usan una gran cantidad de datos. Cientos de empresas emergentes de tecnología financiera están «desagregando» la banca y proporcionando servicios independientes para cada cosa, desde préstamos a asesores robotizados para la gestión de la salud o valoraciones de préstamos. Entre las populares empresas que ofrecen estos servicios están PayPal, Moven, Credit Karma, LendingHome y Wealthfront, por nombrar sólo unas cuantas.

Otro sector de la banca que está en auge es el de los pagos entre particulares. De hecho, cada vez son más los proveedores de tecnología financiera que se dedican a los pagos entre particulares, entre ellos LendingClub, Venmo y PayPal. Más significativo aún es que plataformas sociales como Facebook o SnapChat hayan empezado a allanar el camino hacia los pagos entre particulares. Porque la oportunidad de mercado global de los pagos entre particulares supera con creces el billón de dólares. Y, por supuesto, si hablamos de tecnologías de pagos disruptivas no podemos olvidarnos de Bitcoin. Esta forma de criptomoneda sigue usándose como método de pago, y desde luego es un tema de conversación recurrente en el mundo de la banca y de las finanzas. El índice actual de innovación en los bancos y las

instituciones financieras está aumentando significativamente. Si la IA y los ordenadores pueden aprender a ganarnos en un juego de tablero, ¿por qué no van a ganarnos en cosas donde no podemos permitirnos el error, como los diagnósticos médicos, el análisis de riesgos o el asesoramiento legal y de inversiones? ¿Y qué efecto tendrá esto con el paso del tiempo —pongamos veinte años— en el modo en que se ofrecen diferentes servicios, como los de la banca? La banca no ha cambiado mucho en doscientos años, pero los veinte próximos prometen traer una revolución digital.

La IA ya está entre nosotros y se usa a diario en los pagos, en la gestión financiera y en el asesoramiento robotizado, especialmente en el sector de los asistentes inteligentes digitales que responden a las consultas de los clientes y ejecutan las tareas más habituales. Puede procesar macrodatos con una eficiencia mucho mayor que los humanos, y puede reconocer el habla, imágenes, textos o patrones de comportamiento *on line*, por ejemplo, para detectar fraudes o anuncios idóneos para aumentar las ventas.

El Banco Santander ha anunciado que va a hacer posible que las transacciones a través de su *app* sean seguras gracias al reconocimiento de voz, mientras que el RBS ha puesto a prueba a «Luvo», un servicio de asistencia al cliente para que interactúe con el personal y que en un futuro atienda a los clientes. En sus primeros tres meses, la Nina web del Swedbank sueco ha alcanzado una media de 30.000 conversaciones al mes y un índice de resolución de problemas en el primer contacto del 78 por ciento. Nina puede atender 350 preguntas diferentes de los clientes y dar respuestas. Muchos otros bancos británicos e internacionales han adoptado sistemas similares o los tienen en fase de pruebas.

La tecnología de aprendizaje automático ha avanzado muy rápidamente en los últimos diez años. El ordenador analiza los nuevos datos y los compara con los existentes en busca de patrones, parecidos y diferencias. Repitiendo esta actividad, la máquina mejora su capacidad de predicción y clasificación de información, lo que facilita la toma de decisiones basadas en los datos. Los bancos y las empresas de tecnología financiera ya usan el aprendizaje automático para detectar fraudes, deteniendo las transacciones poco habituales, pero también para otros objetivos. Es más eficiente que el control manual humano y se espera que acabe convirtiéndose en la norma en la banca y las finanzas.

Las marcas que utilizan la IA cultivan una «economía informada» en la que cada movimiento empresarial quede registrado con una precisión máxima gracias al mundo interconectado creado por la internet de las cosas (objetos que envían y reciben datos). Las empresas quieren aprovechar la oportunidad de aplicar la «computación cognitiva», proceso que combina la codificación, las herramientas y los datos para llegar a la inteligencia artificial, a razonar y a aprender. Ahora, más que nunca, los analistas y los científicos de registros de los servicios financieros tienen la capacidad de hacer predicciones concisas basadas puramente en los datos y en el

rendimiento. En particular, éstos y los algoritmos podrían ayudar a mejorar el sistema de calificación de riesgo actual.

Calificar la capacidad de un solicitante para devolver un préstamo no es nada fácil. La capacidad crediticia es un valor de tres dígitos que ofrece a las empresas de préstamo la posibilidad de valorar la seguridad que les ofrecen con respecto al solicitante medio. Independientemente de la percepción que se tenga de una determinada capacidad crediticia, lo que esto quiere decir es que tu capacidad crediticia importa, y que tiene un gran impacto en la facilidad con la que puedas obtener préstamos y con los tipos de interés que se te cobren. En los últimos años han aparecido una serie de empresas emergentes que han redefinido el modo de determinar la capacidad crediticia usando la tecnología *big data* y el aprendizaje automático.

Cuando un joven, un estudiante o un emigrante se muda a un nuevo país o a una nueva ciudad, a menudo tiene dificultades para acceder a un crédito, a un préstamo para el estudio, a una tarjeta de crédito, a un dinero extra para un tratamiento médico o a una pequeña inversión para lanzar un negocio. Desgraciadamente, la mayoría de esas personas no tienen una cuenta corriente en su nuevo lugar de residencia y les cuesta obtener crédito, a pesar de que hayan pagado siempre sus deudas a tiempo y de que accedan a aceptar tipos de interés mucho más altos.

Afortunadamente, la mayoría de estas personas disponen de un teléfono móvil que usan en su vida diaria. Nuria Oliver, Jaime González y José González han demostrado que los móviles y los macrodatos también pueden dar acceso al crédito gracias a los patrones de comportamiento que revelan los datos móviles. Han demostrado que, si los usuarios permiten el análisis de los datos móviles, al cabo de unos minutos puede ofrecérseles un préstamo atractivo, algo que nunca antes habrían imaginado.

Usando técnicas de inteligencia artificial y aplicándolas a los datos de los móviles se pueden crear modelos precisos de capacidad crediticia que puedan mejorar las posibilidades de cientos de millones de personas; personas actualmente *infrabancarizadas*, inmigrantes nuevos en el sistema económico; personas que no tienen acceso al crédito y que, si lo tienen, es con unos intereses excesivos. Los científicos de datos han desarrollado y aplicado nuevos algoritmos de IA para descubrir nuevos patrones de comunicación, movilidad y consumo correlacionados con una responsabilidad financiera.

Por ejemplo, la varianza en el número de llamadas telefónicas, el número de llamadas telefónicas a números particulares o la predictibilidad en el comportamiento y en el radio de la movilidad son algunas de las variables que usa nuestro sistema de IA para crear un modelo de riesgo. Sus algoritmos pueden determinar automáticamente qué variables importan y en qué grado. Por ejemplo, cuanto mayor sea la varianza en el número y la duración de las llamadas realizadas a diario, mayor es el riesgo; pero cuanto más bajo sea el radio de movilidad, más bajo es el riesgo.

Intuitivamente, estas variables detectan aspectos relevantes del carácter de la gente: la gente con una mayor variabilidad en sus patrones de comunicación parecen tener una menor capacidad de previsión de las situaciones futuras y por lo tanto tienden a asumir, de forma inconsciente, riesgos mayores de los que se pueden permitir.

En un estudio piloto realizado en Brasil consiguieron mejorar la fiabilidad de las predicciones de riesgo en un 27 por ciento en comparación con el pionero sistema de valoración creado por una agencia de crédito de treinta años de antigüedad. Eso significa la posibilidad de ofrecer un producto financiero a un 27 por ciento más de personas asumiendo el mismo nivel de riesgo.

Las valoraciones de capacidad crediticia se usan mucho en el sector financiero como modo para controlar el riesgo a la hora de prestar dinero a clientes potenciales. La capacidad crediticia del individuo da información sobre la capacidad de los clientes de devolver sus deudas —o a la posibilidad de que incumplan sus compromisos—. Tradicionalmente, las calificaciones de riesgo se han calculado a partir de la historia financiera del usuario. Las agencias de crédito, como Equifax o Experian, proporcionan esas calificaciones a las partes interesadas combinando toda la información financiera disponible de los clientes con diferentes productos financieros y en diferentes instituciones.

No obstante, la IA puede hacer ese proceso mucho más barato, rápido y efectivo, haciendo llegar el crédito a millones de personas con dificultades para acceder al dinero (por ejemplo, nuevos emigrantes: sólo en Estados Unidos hay 1,1 millones de nuevos emigrantes por trabajo al año, nuevos licenciados universitarios, personas que se recuperan de tropiezos económicos anteriores e infrabancarizados, personas que tienen cuenta en el banco pero que no la usan de forma regular, que en Estados Unidos eran un 20 por ciento de la población en 2013). Éste es un gran ejemplo de cómo usar la IA y los macrodatos de la telefonía móvil para mejorar las posibilidades de la gente hasta límites insospechados.

Según IBM, este año se necesitarán casi 4,4 millones de científicos de datos. El sector inmobiliario no ha sido una excepción en la adopción de los datos como herramienta, incorporando la IA para resolver problemas y predecir el comportamiento de los consumidores. La tecnología *big data* es útil en muchos aspectos, desde el diseño de casas mejores a la mejora del proceso de gestión de proyectos. Por otra parte, los inversores en este sector han encontrado modos de reducir los riesgos mediante el uso de los datos y la IA. Por ejemplo, una serie de empresas emergentes están impulsando el Open Data Movement para mejorar aún más sus análisis del sector inmobiliario, incorporando referencias de criminalidad y de los sistemas de transporte público.

«¿Quieres saber cuál será la gran ocasión para invertir en propiedades dentro de cinco o diez años? Mira dónde viven ahora los artistas», decía un provocador artículo publicado en 2007 en *Business Week* titulado «Bohemian Today, High-Rent

Tomorrow» («Bohemio hoy, rentable mañana»). De hecho da la impresión de que los precios de la vivienda en los diferentes barrios de Manhattan se puede predecir en base a cómo se mueven por la ciudad los bohemios, los artistas y los gays.

Los sociólogos y los políticos siempre han vendido el arte y la cultura como una panacea para las regiones, las ciudades y los barrios deprimidos económicamente. ¿El motivo? Se ha demostrado que los artistas —definidos como artistas visuales autónomos, actores, músicos, escritores, etc.— pueden estimular las economías locales de diferentes modos y elevar los precios de la vivienda urbana.

En muchos casos los artistas son un signo precoz del aburguesamiento de un barrio. Son la vanguardia de lo que se pone de moda. Al tener habitualmente pocos ingresos, suelen verse obligados a buscar barrios baratos en proceso de desarrollo donde puedan pagarse el alquiler. Pero gracias a su creatividad son capaces de mejorar la zona, y acaban por atraer tiendas de moda, galerías y restaurantes. Aunque algunos alcanzan el éxito como escritores, actores, pintores o bailarines, otros se cansan de buscar ingresos trabajando en bares o restaurantes y en ocasiones se lanzan como emprendedores.

También es importante que la vida cueste poco. Pocos artistas en busca de oportunidades pueden permitirse vivir en barrios como el SoHo de Nueva York o Greenwich Village, o incluso en Williamsburg, que en su día fueron grandes centros de actividad artística antes de que atrajeran a residentes más ricos. Ahora es más probable encontrar a los artistas neoyorquinos en el Bronx, en otros barrios de Brooklyn o incluso en Filadelfia.

La tolerancia de una cultura abierta también permite que los artistas atraigan talento y capital humano independientemente de raza, etnia u otras consideraciones. Las poblaciones de artistas se concentran en comunidades que valoran la apertura de mente y la expresión personal. Y su estatus como grupo históricamente marginado hace que las poblaciones de artistas se muestren receptivas a los nuevos llegados, y que apoyen más la innovación, la capacidad emprendedora y la creatividad. Este vecindario bohemio hace que crezcan los precios de la vivienda porque hacen que unos barrios que estaban listos para el crecimiento se vuelvan aún más deseables para un público más amplio.

¿Qué partes de Barcelona podrían ser barrios emergentes? ¿Podría inferirse esta información de los macrodatos? ¿Qué zonas están poblando los jóvenes emprendedores? ¿Qué nacionalidades son las que se mueven de un sitio a otro? ¿Qué estatus socioeconómico tienen? ¿Qué relación mantienen con otros barrios de la ciudad? ¿Determinarán los jóvenes extranjeros residentes en la ciudad o las profesiones más liberales los precios futuros de las propiedades en determinados barrios de Barcelona?

Éste es el mundo de la analítica predictiva de la IA y de los datos masivos; la versión científica de una bola de cristal. En lugar de mirar al cristal miramos una cantidad (idealmente) masiva de aquéllos y recurrimos a técnicas de IA como la estadística, la creación de modelos y el aprendizaje automático buscando patrones que indiquen un comportamiento actual o futuro. Ésa es la historia de una ciudad, la que nos cuenta el alma de sus datos y la inteligencia artificial.

La IA, el empleo y la movilidad

Cuando la gente pierde el trabajo, su comportamiento cambia. Puede que salgan menos de casa, estarán despiertos a horarios diferentes y llamarán por teléfono menos, o a un conjunto de personas diferente. Estos cambios individuales se acumulan en grandes patrones que se pueden observar con la IA y series de macrodatos como las del uso de las redes sociales.

Dos estudios recientes demuestran que la IA y los datos móviles pueden dar pistas sobre nuestro estatus laboral, pueden hacer un seguimiento de la economía y predecir cambios más rápido que con los métodos tradicionales.

Los métodos que usan actualmente los gobiernos para hacer estadísticas macroeconómicas se basan en un paradigma de recopilación y análisis de datos iniciado en los años treinta. La mayoría de estadísticas económicas se construyen a partir de registros de encuestas o de registros gubernamentales.

En la mayoría de países, el Instituto Nacional de Estadística usa diversos métodos para reunir datos estadísticos sobre empleo y desempleo. Estudia las empresas, las familias y las cifras de empleo en el sector privado y en el público, y también recoge referencias mensuales sobre las personas que solicitan ayudas al desempleo. Cada método tiene sus ventajas y sus inconvenientes: por ejemplo, las encuestas a las familias registran datos de individuos de todos los rangos de edad, y de los que no son candidatos a prestaciones de desempleo, mientras que las encuestas a las empresas pueden capturar los de empleo por región y por sector.

Estos análisis múltiples se publican periódicamente, en plazos que van de cada mes y medio a una vez al año. Elaborar y analizar estadísticas lleva mucho tiempo y sale muy caro, lo que significa que a menudo quedan atrasadas y no reflejan la realidad de la economía, aunque compensan el desfase con un gran nivel de detalle y precisión. Aunque seguramente habrá un modo mejor de hacerlo, ¿no?

Detectar el eco del desempleo en los macrodatos resulta mucho más rápido. Como caso práctico, Jameson Toole, Marta González y su equipo del MIT analizaron lo que pasaba con los datos de los teléfonos móviles tras un despido masivo en una fábrica en un lugar de Europa no revelado. En diciembre de 2006 una fábrica de piezas de automoción cerró y 1.100 personas perdieron su empleo, y eso ocurrió en una población de unas 15.000 personas.

Los investigadores usaron datos de servicios móviles para examinar los registros anonimizados de los vecinos de la población. Los investigadores buscaron patrones en el modo en que usaban sus servicios móviles los vecinos antes y después del despido. La gente que trabajaba en la fábrica pero no vivía en la población mostraría una marcada disminución de las llamadas realizadas cerca de la planta tras el día del despido, así que los investigadores se centraron en los individuos que mostraban este patrón. Observaron que estas personas hacían menos llamadas, a un menor número de contactos, y que se reducía su movilidad (el número de antenas por las que pasaban sus llamadas). También analizaron el *churn*, que mide la variación de los destinatarios de las llamadas de un mes a otro. Los desempleados recientes mostraban un *churn* mayor que los sujetos de control, lo que indica que sus redes sociales eran menos estables.

Si es posible identificar una conducta típica de desempleados en los registros de llamadas, también debería ser posible ver esta imagen a una escala mucho mayor que la que da una pequeña población industrial. Para explorar esta posibilidad, los investigadores analizaron a 10 millones de clientes en un país europeo diferente no revelado entre 2006 y 2009. En lugar de buscar despidos en masa, buscaron datos correspondientes a los signos de comportamiento que habían identificado en el estudio de la fábrica.

Usaron esos datos para intentar deducir el índice de desempleo en un trimestre determinado, como el primero de 2007 (T1 2007). También usaron los datos de ese trimestre para predecir los del trimestre siguiente, en este caso el T2 2007. De este modo pudieron comparar sus cálculos y predicciones con las cifras oficiales de estos períodos. Tanto los cálculos del período en curso como los de períodos futuros coincidían mucho con las cifras oficiales, aunque la correlación era algo menor en el caso de las predicciones futuras.

La información de las redes sociales analizada con la IA nos puede decir cosas diferentes sobre cómo se comporta la gente tras quedarse en el paro, pero también puede ser útil. En la Universidad Carlos III de Madrid, Esteban Moro y su equipo usaron 19,6 millones de tuits geolocalizados en España entre 2012 y 2013 para establecer cuántas personas se movían y desde qué hora estaban activas en la red cada mañana, y la corrección gramatical de sus tuits.

Observaron una correlación entre las regiones con un mayor índice de empleo y los tuits que mostraban mayor movilidad, los más madrugadores y los más correctos gramaticalmente. Al igual que en el caso de los datos telefónicos, estos rastros dejados en las redes sociales podrían usarse para hacer cálculos precisos de desempleo en una región determinada.

Esta línea de investigación de momento sólo muestra correlaciones, no una relación de causalidad. Pero da pie a que se hagan más análisis y un mapeo demográfico. Acelerar las predicciones con la IA podría resultar extremadamente útil a la hora de tomar decisiones económicas. Hay una enorme inversión de entidades

públicas y privadas para la medición del estado de la economía de modo que se pueda actuar con mayor conocimiento de causa en las ciudades. Si pudiéramos entender los cambios en el estado de la economía más rápidamente, podríamos intervenir más rápidamente.

Las nuevas *apps* de transporte bajo demanda y de navegación (por ejemplo, Uber o Waze) podrían reducir muchísimo el tráfico de la gente que va en coche al trabajo si los servicios de viajes compartidos y de transporte privado arraigan. Estas *apps* comunican automáticamente la ubicación de los usuarios y su velocidad. Lo mismo ocurre con los coches autónomos o sin conductor. Y los datos de los teléfonos móviles, combinados con los de los mapas, podrían hacer muy efectivos los servicios de viajes compartidos y de conducción sin conductor.

Uber es un servicio de reserva de taxi a través de una *app* que pone en contacto a usuarios que necesitan ir a algún sitio con conductores dispuestos a llevarles. El modelo de negocio de Uber se basa por completo en la IA y en el principio del *crowdsourcing*: cualquiera con coche que quiera ayudar a alguien a llegar a su destino puede ofrecerle un pasaje. Eso da mayores oportunidades a los que viven en zonas donde hay poco transporte público y ayuda a reducir el número de coches en nuestras concurridas calles al compartir vehículo.

Uber almacena y monitoriza los datos de cada viaje que hacen los usuarios y usa la IA para determinar la demanda, distribuir sus recursos y establecer precios. La compañía también hace un análisis a fondo de las redes de transporte público de las ciudades en las que opera, para poder ofrecer cobertura en las zonas de menor servicio y dar conexiones con los autobuses y los trenes.

Uber cuenta con una gran base de datos de conductores en todas las ciudades que cubre, de modo que cuando un pasajero solicita un trayecto, puede ponerle en contacto al instante con los conductores más idóneos. La compañía ha desarrollado algoritmos para monitorizar las condiciones del tráfico y de los trayectos en tiempo real, lo que permite ajustar los precios según cambie la demanda y si las condiciones del tráfico indican que los viajes van a alargarse. Eso anima a más conductores a ponerse al volante en caso necesario —o a quedarse en casa cuando hay poca demanda.

La compañía ha solicitado registrar la patente de este método de tarificación basada en *big data*, que ellos llaman «precios dinámicos». Es una aplicación de la «tarificación dinámica», similar a la que usan las cadenas hoteleras y las aerolíneas para ajustar los precios a la demanda, aunque en realidad en lugar de simplemente aumentar los precios los fines de semana o durante las fiestas supone usar un modelo predictivo para calcular la demanda en tiempo real.

Los viajes compartidos son un modo efectivo de reducir el número de coches en las calles para resolver problemas individuales y de toda la ciudad. Por una parte, los

individuos buscan reducir el coste del uso de su coche y ahorrar en gasolina y otros costes que dependen del uso. Por otra parte, a las ciudades les interesa reducir el tráfico y la polución y ofrecer incentivos (por ejemplo, carriles reservados para los vehículos con más ocupantes) para que la gente comparta vehículo al ir al trabajo. Los sistemas de viajes compartidos nacieron en Estados Unidos durante la segunda guerra mundial. Estos primeros sistemas «boca a boca» requerían quedar con antelación y que los usuarios se conocieran, lo que limitaba el número de vecinos con los que podía compartirse coche.

Los estudios realizados sobre el potencial latente de los viajes compartidos sugieren que podrían tener un impacto significativo en nuestras ciudades y en el tráfico. Los primeros resultados demuestran que compartir coche supone una ventaja económica con respecto a conducir solo, y que es una solución más flexible y que consume menos tiempo que el transporte público, pero no está claro que estas ventajas tengan el peso suficiente como para animar a los trabajadores a pasarse a los viajes compartidos; la intimidad y la flexibilidad son los principales motivos que esgrimen la mayoría de los que prefieren ir solos en su coche al trabajo.

Muchos creen que la tecnología actual ofrece un nivel de seguridad y seguimiento insuficiente como para que la gente pueda viajar tranquila con extraños; otros creen que es sólo un problema de falta de iniciativa que impide que la tecnología se desarrolle. La mayoría de personas, no obstante, suponen de un modo implícito que los patrones de movilidad humana de las ciudades y su trazado actual presentan suficiente espacio como para que los viajes compartidos arranquen, en cuanto se resuelvan problemas como los mencionados.

En un estudio de la Universidad de California en Irvine y Telefónica se usaron datos anonimizados de movilidad y de Twitter para calcular que el tráfico de la ciudad de Nueva York podría reducirse en un 20 por ciento si la gente estuviera dispuesta a compartir coche con extraños que viven y trabajan a menos de un kilómetro de sus lugares de residencia y de trabajo, y si toleraran perder un máximo de diez minutos más en los traslados. En Los Ángeles, la cifra equivalente fue del 13 por ciento, menor debido a la densidad de población de la ciudad, mucho más baja. Este modelo sugiere que, si la gente estuviera dispuesta a recoger y dejar pasajeros durante el recorrido, el tráfico de Nueva York podría reducirse en un 44 por ciento, y el de Los Ángeles en un 40 por ciento.

Google parece dispuesta a desafiar a Uber adelantando el uso de la tecnología de los coches sin conductor y a ofrecer un servicio de viajes compartidos. El éxito de cualquier interfaz de viajes sin conductor depende de lo bien que procese el ordenador del coche los datos que envía al sistema de inteligencia artificial. Los sensores LIDAR, GPS, altímetros, giroscopios y taquímetros instalados en el coche transmiten enormes cantidades de datos al programa de *deep learning* del ordenador de abordo, que usa esta información para alterar los algoritmos y dirigir los

movimientos del coche en tiempo real, algo que tiene que hacer a una velocidad y con una precisión increíbles.

A largo plazo, los mejores coches sin conductor serán los que tengan la mejor IA, que sean capaces de procesar datos de navegación rápidamente y con precisión. Para que funcione y valga la pena, la IA tiene que ser más lista que los miles de millones de humanos que matan a 1,2 millones de personas cada año con sus coches, y con más datos y mejores algoritmos informáticos enseguida podremos hacer progresos. Con resultados así, podríamos empezar a ver muchos menos coches en las ciudades, y más gente compartiendo vehículo con otros que cubren un trayecto similar, ahorrando dinero y aumentando la seguridad de las calles.

6

La IA en busca del bien global

La medicina ha sido una ciencia de los médicos con apoyo de datos. La medicina está a punto de ser una ciencia de datos con apoyo de los médicos.

CRAIG VENTER

Un día en Harvard

Es un día fresco y soleado en Boston. Estamos cerca de acabar una mañana de reuniones en el MIT y en Harvard. Yo, en particular, tengo curiosidad por conocer al profesor Clay Christiansen, que ha dado con una teoría de innovación disruptiva, esto es, de innovaciones que crean negocios completamente nuevos, no sólo que mejoran negocios existentes (por ejemplo, la llegada de la industria de las neveras y la congelación frente a la de conserva de alimentos en sal). Yo soy un gran defensor de construir modelos de cómo funcionan las cosas, ya que eso puede ayudarte a razonar y a hacer grandes innovaciones, pero también estoy convencido de que todos los modelos son claramente erróneos.

Dado que las cosas cambian constantemente, todos los modelos pasan por alto una parte de la realidad, y crear un modelo innovador probablemente sea muy difícil. No obstante, la diferencia entre los buenos y los malos es que algunos son más útiles que otros para comprender ciertas cosas. Tengo mucha curiosidad por oír hablar de sus teorías y modelos sobre cómo innovar con ambición y crear nuevas disrupciones, entre las cuales cuenta con recetas para crear una nueva empresa independiente que innove en proyectos muy ambiciosos de gran impacto social.

En los últimos años se ha hablado mucho del concepto de la «innovación disruptiva». El mismo profesor Clayton M. Christensen incluso ha moderado la euforia, aclarando lo que quiere decir al usar este término. A pesar de las grandes diferencias en la aplicación que se hace de ella, la mayoría coincide en que las innovaciones disruptivas:

1. Son más baratas (desde el punto de vista del cliente).
2. Son más accesibles (desde una perspectiva de usabilidad o distribución).
3. Usan un modelo de negocio con ventajas estructurales de precio (en relación con las soluciones preexistentes).

Si estas características de la disrupción son importantes es porque cuando las tres están presentes es difícil que una empresa existente pueda plantear competencia.

Para la reunión hemos quedado en el Harvard Faculty Club, edificio emblemático del campus de Harvard inaugurado en 1931 y en el que se exige que los hombres lleven americana y corbata. Nos reciben con canapés fríos y una copa de cava. Entre los invitados hay emprendedores de Boston, profesores y excolegas del MIT con los que he realizado investigaciones sobre redes.

En el momento en que me acerco a la barra en busca de otro vaso de agua con gas se me cruza el doctor Rosenquist, que se presenta con gran elegancia y elocuencia. Ha venido a verme desde el General Massachusetts Hospital, nombrado el mejor hospital de Estados Unidos y donde se inventó y se usó por primera vez la anestesia, en 1846. Lo que más le interesa es la posibilidad de conectar la tecnología, los datos masivos y los sensores para conseguir que la gente viva más tiempo con buena salud, y en particular la aplicación de alguna de las teorías del professor Clay Christiansen para reinventar el ámbito sanitario.

Intercambiamos tarjetas de visita y quedamos para el día siguiente en las instalaciones del Media Lab en el campus del MIT. Tengo la impresión de que es una ocasión más para sacar la tecnología del mundo de la tecnología y conseguir que tenga un impacto en el mundo real, en algo que afecta al día a día de la gente. Así que espero el día siguiente con esperanza, listo para embarcarme en una nueva aventura.

Resultó que el doctor Rosenquist era un experto psiquiatra muy preocupado por las enfermedades mentales. Quedamos para almorzar en un café cercano, junto al Charles River, con vistas del *skyline* de Boston, entre dos ciudades de lo más sugerentes, Boston y Cambridge —la de Massachusetts, sede del MIT y de Harvard—. Su esposa es psicóloga infantil, y dado que ambos teníamos un hijo prácticamente de la misma edad, lo primero de lo que hablamos fue de las cosas importantes de la vida (por ejemplo, de cómo conseguir que los niños se duerman).

Me sugirió que leyera un libro del doctor Richard Ferber, que básicamente describe el «método Ferber», basado en rituales nocturnos que combinan intervalos de tranquilizar al niño con otros en que se le deja solo, lo que a la larga debería inducir un comportamiento con el que los bebés se acostumbraran a dormir toda la noche.

Más tarde descubriría que existen numerosas aplicaciones para ayudarte a aplicar este método. Y éste no es más que un tipo de las muchísimas *apps* de salud y bienestar que van encontrando hueco en nuestra vida diaria. ¿Y si la próxima vez que vayas a la farmacia o al psicólogo te aconsejaran una *app*, o si el médico de familia te prescribiera una combinación de *apps* que usaran la tecnología *big data* y la IA para cuidar la salud y la forma física, en lugar de seguir una dieta?

Por ejemplo, hoy en día los padres y madres pueden hacer un seguimiento del crecimiento y el desarrollo de su bebé prematuro con sencillas aplicaciones. Los

padres introducen la información a diario con una tableta. Esa información incluye datos sobre signos vitales, lo que ha comido el bebé o el número de pañales que ha mojado o manchado. Los datos se evalúan mediante algoritmos de IA que activan alertas al equipo de enfermería, lo que puede llegar a salvar la vida al bebé. Otras *apps* permiten interactuar con los hijos a distancia mediante la realidad virtual. Padres e hijos entran en el mismo mundo virtual, donde pueden jugar e interactuar. Por ejemplo, el padre o la madre puede narrar una historia que haga que ambos viajen a lugares mágicos sin salir de la cama.

Hoy en día los padres compran clips inteligentes para pañales, peleles inteligentes, calcetines inteligentes y tobilleras inteligentes para evitar que los bebés mueran durmiendo. Las muertes durante el sueño, que incluyen el síndrome de muerte súbita del lactante, o SMSL, matan a unos 3.500 bebés al año sólo en Estados Unidos También son difíciles de detectar —aparte de comprobar que el bebé hinche y deshinch el pecho, los padres tienen muy pocos recursos—. Aunque los estudios sobre si la monitorización de los bebés reduce el riesgo de SMSL aún no son conclusivos, cabría esperar muchos adelantos en este campo combinando sensores y datos.

La IA y la salud mental

Al doctor Rosenquist le interesaba saber cómo podían ayudar los macrodatos y la IA en particular a la salud mental. Está convencido de que la tecnología móvil revolucionará la asistencia sanitaria, empezando por la salud mental. Obtener datos inmediatos relacionados con la salud mental del paciente (estado de ánimo, etc.) es de importancia capital, y los nuevos biomarcadores electrónicos para la salud mental serán como un ECG de la salud mental. Para llegar a esto, hay que establecer relaciones entre ingenieros informáticos, científicos de datos, médicos y hospitales/proveedores de servicios sanitarios que estén familiarizados con las necesidades y los métodos de los otros.

Rosenquist me explicó que ciertas enfermedades mentales, que se miden a través de los recuerdos del paciente —en muchos casos precisos— en una entrevista, se basan en una serie de síntomas que pueden controlarse con una combinación de *big data* e IA, innovadores sensores y teléfonos móviles. Sugirió que actualmente los médicos poseen de una imagen poco precisa de la evolución del paciente, basada únicamente en los síntomas del paciente durante las visitas, pero que puede mejorarse mucho con los datos recogidos con dispositivos móviles, que ofrecen una información mucho más detallada, a una escala nunca vista antes, y mucho mejor que la obtenida con estudios intermitentes sobre el estado de la persona. Por ejemplo, los patrones del sueño, los niveles de energía, los niveles de concentración, los síntomas psicomotores (levantarse de la cama) o el interés en diferentes actividades son

síntomas de trastornos psiquiátricos que se pueden medir fácilmente a través de dispositivos móviles.

Quedamos en vernos de nuevo en Barcelona y desarrollar esas ideas aún más. Yo volví a casa con ganas de probar el método Felber y esperando que así todo el mundo pudiera dormir mejor por la noche, pero seguía pensando en la relación entre *big data*, IA, teléfonos móviles y salud.

Hace cinco años, Ben Lipkowitz, que ahora tiene veintiocho años, vivía con unos amigos en Bloomington (Indiana, Estados Unidos), y se le ocurrió preguntarse cuánto tiempo se pasaba lavándole los platos a uno de sus compañeros de piso. Lipkowitz tenía un *datebook* electrónico que se había comprado en un viaje a Tokio, y el 11 de mayo de 2005, a las 14.20, empezó a usarlo para llevar un registro de sus acciones. En lugar de introducir sus citas futuras, introdujo sus actividades pasadas, creando un relato de su vida bastante completo.

En cierto sentido no era más que un diario personal normal y corriente, sólo que en formato digital y con un detalle poco habitual. Pero el formato y el detalle eran lo que marcaba la diferencia. Lipkowitz posteriormente transfirió todos los datos a su ordenador, y ahora, con unos cuantos comandos abreviados, puede recuperar su historial. Sabe cuánto ha comido y cuánto ha gastado. Sabe qué libros ha leído y qué objetos ha comprado. Y, por supuesto, tiene la respuesta a su pregunta de partida: «Yo pensaba que me pasaba una hora al día lavando lo que ensuciaba esta persona —dijo Lipkowitz—. Y resultó que eran más bien veinte minutos».

Usamos números cuando queremos revisar un coche, analizar una reacción química, predecir el resultado de unas elecciones. Usamos los números para optimizar una línea de montaje. ¿Por qué no usar los números y los datos con nosotros mismos?

El contraste con la noción terapéutica tradicional de desarrollo personal es sorprendente. Cuando nos medimos a nosotros mismos, no tenemos el imperativo de mirar más allá de nuestra existencia diaria y observar una verdad oculta a un nivel más profundo, sino que entendemos la personalidad de nuestros pensamientos y acciones más triviales, la personalidad que, sin ayuda técnica, podría pasarnos desapercibida, como la personalidad que deberíamos llegar a conocer. Por encima de la atracción que generan las cifras de los macrodatos de uno mismo está la suposición de que muchos de nuestros problemas proceden simplemente de la falta de instrumentos para comprender quienes somos.

Nuestros recuerdos son pobres; somos vulnerables a muchos tipos de sesgo; sólo conseguimos fijar la atención en una o dos cosas a la vez. No tenemos un podómetro pegado al pie, ni un espirómetro en los pulmones, ni un monitor de glucosa instalado en las venas. Nos faltan las herramientas físicas y mentales necesarias para evaluarnos a nosotros mismos. Necesitamos ayuda de las máquinas y de la IA.

Según *The Economist*, hace cien años un audaz investigador fascinado por el misterio de la personalidad humana haría uso de nuevos conceptos psicoanalíticos como la represión y el inconsciente. Estas ideas fueron inventadas por personas a las que les encantaba el lenguaje. Y aunque los conceptos terapéuticos de la naturaleza personal iban haciéndose cada vez más comunes y adquirirían una forma simplificada y más accesible, conservaban algo del elaborado humanismo literario de sus inventores. Desde la languidez del diván del psicoanalista a la indiscreción de un cuestionario de autoayuda, las formas dominantes de la exploración personal suponen que el camino hacia el conocimiento se fundamenta en las palabras.

Ahora podemos explorar una ruta alternativa. En lugar de investigar los mundos interiores de la gente a través del habla y la escritura, podemos usar la IA y las cifras de los macrodatos. Aún estamos lejos de retirar de escena para siempre el diván del psiquiatra, pero el uso de la IA en la exploración personal promete mucho, aunque sólo sea porque ofrece una respuesta a los que a la pregunta «¿Quieres hablar de ello?» responden con un sonoro «No».

Hace poco vi un mensaje de correo electrónico de un director de cine de veintiséis años llamado Toli Galanis que sigue unos cincuenta flujos de datos personales diferentes, que incluyen actividades, salud, las películas que ha visto y los libros que ha leído, los amigos con los que habla y los temas de los que tratan. Aunque Galanis ha reconocido que disfruta recopilando datos y organizándolos de forma inteligente, lo que me llamó la atención fue otro aspecto de su declaración. «Sé que inmediatamente después de ver una película mala tiendo a ser más negativo con respecto a mis perspectivas profesionales como director de cine», escribió, a lo que añadía que desde que hacía ese seguimiento se le daba mejor detectar la influencia de circunstancias aparentemente triviales sobre su estado de ánimo y sus decisiones.

La idea de que nuestra vida mental se ve afectada por causas ocultas es un pilar de la psicología. La facilidad para gestionar el flujo de pensamientos y emociones es señal de felicidad y de equilibrio. Pero ¿cómo se hace? Prácticamente todas las prescripciones terapéuticas implican una invitación a observar, a prestar atención. Una vez tenemos una noción en el punto de mira, podemos atacarla con un arsenal de herramientas: cognitivas, psicoanalíticas o incluso espirituales. Pero ninguna de ellas nos dirán si se nos ha pasado algo por alto. Puede que simplemente se te haya pasado por alto un hábito debilitante, una correlación negativa o una mala influencia.

La observación de Galanis al darse cuenta de que las películas malas le provocaban un desánimo en lo profesional es el tipo de conclusión que parece estar al alcance de cualquiera con una mínima conciencia personal; descubrirlo no resulta más difícil que ver un dólar tirado en la calle y recogerlo. Pero por cada dólar que coges, ¿cuántos se te pasan por alto?

Por supuesto, a veces se nos pasa por alto lo que hacemos porque estamos motivados para no verlo. Nos avergonzamos de nosotros mismos, así que nos mentimos. Shaum Rance empezó a hacer un seguimiento de su ingesta de alcohol hace dos años, después de que a su padre le diagnosticaran una enfermedad hepática terminal. No decidió dejar de beber; no hizo un minucioso inventario moral; simplemente se puso a contar, usando la web anónima drinkingdiary.com. Y observó que su memoria externa tenía un gran poder.

Con este registro de datos de las copas que tomaba, adquirió más conciencia y aumentó la sensación de autocontrol, con lo que acabó bebiendo menos. Como la cuenta la lleva una máquina, no siente la vergüenza social que habría podido hacerle infravalorar, conscientemente o no, la ingesta de alcohol. «Al diario no le miento», explica. Al fin y al cabo, sería una tontería fingir ante una máquina y esconderse ante los resultados. El sistema de control es una extensión de una facultad básica de la conciencia de Rance, que le recuerda la realidad, y que lo hace sin ninguna emoción de por medio. Y tal como lo ve él, eso es una virtud.

Recoger estas débiles señales de IA nos da perspectiva. Margaret Morris, psicóloga clínica e investigadora de Intel, hizo recientemente una serie de ensayos de campo usando el teléfono móvil para llevar un registro de emociones. En momentos al azar, el teléfono sonaba e interrogaba a su propietario sobre su estado de ánimo. Un participante de uno de los estudios de Morris repasó las tendencias que mostraban sus datos y observó que cada día se ponía de mal humor a la misma hora. Efectuaba una transición acelerada del trabajo al hogar. Aunque aún tenía en mente las tareas inacabadas, introducía nuevas demandas, y el estrés le acompañaba el resto de la tarde. Los datos le demostraron cuál era el problema. Con esta ayuda aprendió a hacer una breve pausa mental en aquel momento y encontró un gran alivio.

Los psiquiatras buscan sofisticadas herramientas computacionales y de IA que puedan desenmarañar las complejidades de las enfermedades mentales y mejorar las decisiones de tratamiento. Se ha investigado y experimentado mucho para intentar comprender algunas enfermedades mentales incurables. A pesar de que se han encontrado modos de aliviar los síntomas, la prevención y el diagnóstico precoz siguen siendo un problema en la búsqueda de una cura potencial. La detección precoz de los cambios cognitivos puede ralentizar significativamente el desarrollo de la enfermedad, sus complicaciones, y evitar sus devastadoras consecuencias. La inteligencia artificial, basada en la ciencia cognitiva y con su capacidad para procesar cantidades de datos ingentes, puede ser ese elemento que nos faltaba para resolver el misterio que envuelve los trastornos neuronales.

La IA ofrece una enorme cantidad de oportunidades para mejorar la salud y la asistencia social. Y cada vez se presta más atención al valor que puede tener en la

mejora exponencial de la investigación y el diagnóstico de las enfermedades mentales en particular.

Pasamos muchos años intentando diagnosticar enfermedades mentales complejas y variadas, y en identificar posibles tratamientos para los pacientes. Parte de la dificultad que experimentamos radica en que no hay una relación clara entre un diagnóstico, sus síntomas y la causa biológica. Dos pacientes podrían tener la misma afección y los mismos síntomas —pongamos, por ejemplo, una depresión—, pero el origen de la afección podría deberse a dos causas diferentes dependiendo de una variedad de factores de cada individuo, biológicos o de otro tipo. Del mismo modo, dos personas con la misma afección podrían no compartir ningún síntoma, e incluso podría pasar que dos personas tuvieran los mismos síntomas y enfermedades diferentes.

La complejidad de todas las variables implicadas crea graves problemas a los psiquiatras para identificar la causa de las diversas enfermedades mentales, por lo que necesitamos herramientas más potentes para establecer relaciones entre síntomas, trastornos y causas. Ahí es donde pueden ser de ayuda la IA y los análisis de datos.

A la hora de estudiar la salud mental y el cerebro, los neurólogos tienen acceso a unos conjuntos de datos significativos —neuronal, anatómico, de actividad, cognitivos, clínicos y genéticos, entre otros— que se recopilan durante la investigación y el tratamiento.

Hasta ahora no teníamos dónde almacenar estas grandes cantidades de datos, y mucho menos podíamos aprovechar todo su valor, las oportunidades y la información que nos daban para sustentar e impulsar nuevas vías de investigación de las enfermedades mentales.

No obstante, con la incorporación de la computación en la nube, el desarrollo de la IA y de las herramientas de análisis de datos, el panorama ha cambiado.

Los análisis avanzados realizados con la información recopilada a partir de diferentes fuentes de datos permiten secuenciar el genoma y elaborar un mapa cerebral, lo que nos da la oportunidad de descubrir genes y zonas del cerebro hasta antes desconocidos. Esta nueva información tiene un valor enorme y puede ampliar los horizontes de nuestros conocimientos actuales y de nuestra comprensión de la salud mental.

Combinando diversas fuentes de datos de pacientes, y utilizando técnicas de aprendizaje automático para detectar patrones relevantes, conseguiremos entender mejor la relación entre síntomas, diagnóstico y tratamientos, lo que nos permitirá mejorar las decisiones clínicas que tomemos en el futuro, identificar los tratamientos más efectivos para determinadas situaciones y detectar trastornos de forma precoz.

La posibilidad de aplicar la IA y recopilar datos de equipamiento y *apps* es una gran ventaja para los investigadores y los psicólogos. El tratamiento de una persona con depresión solía suponer interminables interrogatorios, a veces incluso semanales. «¿Hoy está contento?» «¿Ayer sentía angustia?» Eso provoca numerosos problemas:

en primer lugar el sujeto quizá no sepa cómo informar, reconocer o cuantificar sus propias sensaciones; en segundo lugar, este enfoque pasa por alto que la depresión, como otros estados, no se mantiene estática ni presenta siempre el mismo aspecto; y además es una simplificación de lo que pasa por la mente o por la vida de la persona.

La posibilidad de acceder a los datos de las redes sociales, de los *smartphones* o de equipos portátiles para su análisis, e incluso de compararlos con conjuntos de datos masivos, podría llevarnos a un modo de interacción completamente nuevo. Ciertas conductas en las redes sociales se han asociado incluso con la depresión posparto con admirable precisión.

Lo que realmente supone un cambio es la nueva generación de asombrosos algoritmos de IA. Estos algoritmos no sólo valoran los cambios de estado y los textos «sobre mí», sino que también analizan las listas de amigos, establecen correlaciones y detectan densidades —lo fuerte que es realmente una red social—. Analizan cuándo se ha unido el usuario al mundo de las redes sociales, la frecuencia de sus *posts*, y las páginas que dice que le gustan o las *apps* que usa. Estos detalles pueden convertirse en una información muy específica sobre la personalidad de un individuo.

A los tutores universitarios se les asignan tareas ingentes, desde ayudar a los estudiantes en su camino hacia la graduación hasta el tratamiento de sus trastornos mentales. Estas tareas pueden convertirse en algo difícil de gestionar cuando cada tutor tiene que ocuparse de muchos estudiantes, sobre todo por lo limitado de sus herramientas.

En muchas universidades, una de las herramientas que se usan para rebajar la carga de los tutores es el análisis predictivo. La colección de datos de los estudiantes, sean informatizados o en papel, siempre han suscitado preocupaciones sobre la privacidad del estudiante, la marginalización de ciertos grupos de estudiantes, e incluso la falta de transparencia. Todos estos factores tienden a magnificarse cuanto más crece el grupo.

Sin embargo, en la Georgia State University han encontrado un modo de usar la IA con sus 24.000 estudiantes para mejorar diversos factores. La universidad, en colaboración con una consultora externa, analizó los datos de 2,5 millones de notas obtenidas por los estudiantes a lo largo de siete años. Este análisis produjo una lista de factores que influyeron en la posibilidad de cada estudiante de llegar a graduarse, y creó un sistema que enviaba señales de aviso precoces, no sólo a los estudiantes, sino también a sus asesores. El sistema usa más de 700 alarmas para identificar los problemas del estudiante para cumplir con los requisitos necesarios para completar sus estudios, y para ayudar a los profesores a actuar en consecuencia.

El progreso de la tecnología informática, de los procedimientos estadísticos y el uso de la IA para la predicción del comportamiento humano están arrojando luz sobre

asuntos tan importantes como la salud mental, la educación, las relaciones íntimas y muchos otros.

La salud física y la IA

De vuelta en Asturias, donde abunda la comida fresca y las maravillas del mar Cantábrico se combinan con las delicias de los Picos de Europa para crear una combinación de experiencias dignas del paraíso en la Tierra, fui a ver al doctor Carlos López Otín, científico de fama mundial que trabaja en el genoma del cáncer. Quedé con él en Oviedo, donde tiene su despacho, con una mesa de escritorio cubierta de estudios y registros científicos que casi impiden verlo cuando se sienta en su silla.

El doctor López Otín se toma su tiempo para describir su investigación, la aplicación del análisis de datos para la descodificación del ADN; coge una hoja de papel y me explica con elocuencia cómo intenta descodificar el ADN de determinados tumores. Recientemente ha descodificado el genoma de la leucemia linfocítica crónica. Los avances en la biología molecular del cáncer realizados en las últimas décadas han determinado que es una enfermedad causada por la acumulación de daños genéticos en las células normales, pero hasta ahora la identificación de estos cambios está resultando ser un proceso lento y laborioso.

No obstante, gracias a los equipos de secuenciación genética más modernos, este proceso se ha acelerado, y ahora mismo este centro puede secuenciar hasta seis genomas humanos en un día. Para analizar el volumen de datos generados en este proyecto han sido necesarios programas especializados. La herramienta de datos y software desarrollada por la Universidad de Oviedo para este objetivo se llama Sindrón, y ha resultado ser esencial para la identificación de las mutaciones presentes en los genomas de los tumores. Con más datos y IA el proceso solo se acelerará y abaratará aún más. Hoy ya hay secuenciadores de ADN en formato USB por menos de 1.000 Euros (MiniON), reduciendo de manera considerable los tiempos y costes.

El doctor Otín me cuenta que el cáncer es algo que todos acabaremos sufriendo si vivimos lo suficiente, pero también menciona algo igual de sorprendente: que aunque hoy en día dedicamos el 80 por ciento de nuestros esfuerzos a solucionar problemas relacionados con enfermedades crónicas, apenas nos fijamos en los patrones de conducta que podrían señalar afecciones graves de forma precoz. Parece que aquí hay un campo en el que trabajar, y donde la IA y los datos genéticos y epigenéticos pueden ayudar mucho a predecir patrones nocivos. Hoy una gota de tu sangre puede predecir una foto de tu cara, lo que no debería ser sorprendente porque los gemelos comparten el mismo ADN. Con IA, 1.100 muestras de genomas y 1.100 fotos, e identificando el gen correspondiente al tamaño de la nariz, color de los ojos, etc. una muestra de ADN puede usarse para crear imágenes o avatares que se parecen mucho a ti o a un familiar.

Craig Venter, el biólogo estadounidense que arrancó su propio proyecto Genoma Humano en 1999, acaba de lanzar sus propios hospitales (Health Nucleus) para detectar enfermedades en estadios tempranos o antes de que pasen (cánceres, aneurismas, problemas neurológicos). Esto lo consigue a través de una combinación de escáneres, análisis de sangre, estudios del corazón y sobre todo con el análisis de tu ADN y de 150 gigabytes de datos obtenidos de tu cuerpo que con la ayuda de IA consiguen en tres horas y por varios miles de Euros hacer un chequeo exhaustivo y preventivo de tu salud.

Estoy en Asturias sólo unos días, para ver a mi familia antes de volar a Nueva York, donde voy a encontrarme con «el hombre más conectado del mundo». Chris Dancy, conocido como el hombre más conectado del mundo, tiene cuarenta y cinco años y entre 300 y 700 sistemas activos en todo momento, sistemas que capturan datos sobre su vida en tiempo real. Sus muñecas están cubiertas con tecnología portátil de diversa índole, como el monitor de actividad Fitbit y el reloj inteligente Pebble. Se pesa en la báscula wifi Aria, en casa usa la iluminación Hue, controlada por teléfono, y duerme sobre una funda de colchón Beddit para hacer un seguimiento del sueño. Incluso sus perros llevan el controlador Tagg, que registra sus actividades diarias.

Dispone de un registro de todos sus datos médicos (radiografías, análisis de sangre, informes, etc.) de hospitales, consultas y laboratorios médicos, todo ello escaneado y procesado digitalmente. Cuando va al médico, lleva consigo terabytes de datos para que éste pueda analizarlos. Bromea diciendo que ningún médico quiere hacerse cargo de su salud. No me extraña... Probablemente él le haga más preguntas al médico de las que éste pueda responder. Pero está decidido a usar la IA para cuidar la salud. De hecho, antes tenía sobrepeso, y ha conseguido perder 45 kilos.

Almorzamos juntos y aprovechando la ocasión, le pregunto sobre la dieta que siguió para perder todo ese peso. Me esperaba que me hablara de una combinación mágica de dieta que le hubiera hecho pasar hambre, o de algún tipo de intervención gástrica para adelgazar. Pero su respuesta es sorprendentemente reveladora.

La principal diferencia fue desarrollar una percepción de su situación en cada momento, en términos de ingesta calórica, y en darse cuenta de cuáles eran las situaciones nocivas que le hacían comer de más. Por ejemplo, darse cuenta de que había sido un día de mucho trabajo y de que su tendencia natural al llegar cansado a casa sería ir a la nevera y echar mano del helado y de una cerveza.

De este modo podía prepararse para esas situaciones y por ejemplo colocar la cerveza en otra parte de la casa, para evitar lanzarse a esa espiral negativa de combinar cerveza y helado, lo que suponía una gran ingesta calórica con un valor nutricional muy bajo.

Cuando empezó a realizar un seguimiento de sus hábitos, examinó atentamente los días en que aumentaba su ingesta calórica, y vio que en muchos casos dependía de con quién estaba o en qué restaurantes (frente a cuando comía en casa). Dancy empezó a evitar comer con determinadas personas, a ciertas horas y en ciertos lugares, y empezó a quedar con ellos sólo para tomar café, y estableció filtros para conseguirlo. Perdió 15 kilos sólo eliminando los desencadenantes que le condicionaban. Era mucho más consciente de cómo responder a la vida y tomar medidas para ajustarse a su entorno. Se creó hábitos más sanos gracias al *feedback* de los datos que obtenía. Con el seguimiento de sus hábitos diarios y la introducción de pequeños cambios consiguió mejorar su salud.

Desde el nacimiento de mi hijo yo también había acumulado unos cuantos kilos. Cambiar de alimentación o hacer ejercicio no estaba dándome resultados, y el cambio de mis patrones de vida estaba empeorando aún más las cosas. No es que me preocupara excesivamente, y no había hecho dieta nunca, pero la charla con Chris Dancy me animó a experimentar y ver qué efecto podían tener los datos en mí mismo. Así que me instalé una *app* para móvil que me ayudaba a contar el número de calorías que tomaba a diario. Esta *app* la usan actualmente 100 millones de personas, e incluye una base de datos exhaustiva de ingredientes y platos, y también cartas de restaurante con un listado de calorías para cada artículo.

La *app* me preguntaba a varias horas del día qué estaba comiendo (por ejemplo, zumo de naranja, bocadillo de jamón, ensalada de tomate, etc.), y la cantidad (por ejemplo, una ración o media, o incluso el peso en gramos). Antes de empezar, la *app* también me había preguntado mi peso de partida, la altura y el peso deseado.

Yo introduje escrupulosamente los datos de todo lo que comí durante tres semanas, suficiente como para tener una buena muestra, durante el desayuno, el almuerzo y la cena, incluidos líquidos y tentempiés. Y un mes después de haber iniciado mi análisis obtuve un resultado inesperado: ¡El análisis de datos me decía que estaba comiendo mucho menos de lo que debía! Unas 500 calorías menos al día, lo que significaba que estaba muy por debajo de mi índice metabólico basal, que es el número de calorías necesario para mantener el cuerpo vivo y funcionando, aunque te pases el día tirado en la cama y no muevas un músculo. Visité a un nutricionista, y me dijo que tenía sentido. Cuando el cuerpo no cuenta con las suficientes calorías, se pone en modo de ahorro de calorías, y todo lo que comes lo acumula en forma de grasa, en preparación de posibles períodos de hambre futuros.

Para perder peso es necesario reducir el consumo de calorías. Pero si reduces demasiado el consumo de calorías, el metabolismo se vuelve tan lento que tu cuerpo entra en un estado en el que la pérdida de peso se detiene por completo. Algunas personas también creen que el no comer suficientes calorías no sólo impide perder peso, sino que puede provocar que el peso aumente. Así que básicamente comer demasiado poco impide el adelgazamiento. En algunos casos incluso puede provocar

ganar peso. Para «salir» de ese estado y empezar a perder peso hay que comer más calorías, no menos.

Si las cifras eran correctas, ¿suponía que tenía que comer más para perder peso? La *app* sugería que así era. Supongo que no tenía nada que perder. En el peor de los casos, ganaría un par de kilos más y luego pararía. ¡Pero funcionó! Durante unos dos meses comí la cantidad de calorías que recomendaba la *app*, lo cual en algunas ocasiones significaba obligarme a comer más. ¡Y *voilà!* Perdí 1,5 kilos al mes durante 4-5 meses hasta que recuperé el peso que me había propuesto como objetivo.

Desde entonces no he vuelto a ganar peso, sobre todo porque ahora sé qué es lo que me funciona, cuáles son los estímulos que hacen que opte por determinados alimentos que me convienen y cuáles evitar; a qué horas del día necesita ingerir calorías mi cuerpo y a qué horas no; cómo mantener los niveles de energía y la ingesta de calorías a un nivel suficiente y constante para no tener ataques de hambre o picos de azúcar que el cuerpo no puede procesar con la suficiente velocidad y que acaba convirtiendo en grasa. De hecho también me he vuelto más consciente de mí mismo gracias a los macrodatos. Ahora tengo una visión de mí mismo a través de los *big data* que espero que me sirva durante muchos años para mantenerme sano y en forma.

Hay decenas —si no cientos— de *apps* de adelgazamiento en el mercado, y suelen clasificarse en dos categorías: alimentación y fitness. Las *apps* de tipo diario alimentario te ayudan a llevar un seguimiento de lo que comes y de las calorías, las grasas, los azúcares, los carbohidratos, las proteínas, etc. Muchas de estas *apps* requieren la introducción manual de la información sobre los alimentos que has comido, que luego se comparan con la base de datos para calcular la información nutricional de los alimentos. Las *apps* de fitness te ayudan a hacer un seguimiento del ejercicio que haces e incluso te proporcionan programas físicos y sugerencias. Entre los programas más populares están el «*Couch to 5K*», los entrenamientos de fuerza, de yoga, y el «entrenamiento científico de 7 minutos».

Las básculas inteligentes también están ganando popularidad. En lugar de limitarse a dar el peso, pueden identificar a la persona y diferenciar a los miembros de la familia. Pueden enviarte los datos del peso al teléfono o al ordenador, o sincronizarse con una *app*. También pueden calcular el porcentaje de grasa corporal, hacer un seguimiento de tus progresos y darte mensajes de ánimo.

Llega un momento en que los propósitos de Año Nuevo empiezan a fallar, y en que los objetivos de adelgazamiento van quedando aparcados, salvo en el caso de los más decididos. Así que la empresa de prendas y accesorios de deporte Under Armour decidió estudiar qué era lo que hacían de un modo diferente los que lo conseguían, estudiando los datos recogidos con su *app* MyFitnessPal.

En un estudio realizado con 4,2 millones de usuarios se observó que 427.000 están a menos de un 5 por ciento de su objetivo de adelgazamiento; Under Armour intentó descubrir cuál era su secreto. El resultado fue sorprendente. Los usuarios que lo conseguían no comían menos carbohidratos, azúcares o grasas; de hecho su ingesta de calorías era más o menos la misma que los que tantas dificultades tenían para seguir su dieta.

¿Cuál era la gran diferencia? Que ingerían mucha más fibra. Los que consiguen su objetivo de adelgazamiento comen un 29 por ciento más de fibra, una media de 13,5 gramos frente a los 10,5 g de los que no alcanzan sus objetivos.

Curiosamente, hasta esos 13,5 g están significativamente por debajo de los 25 g o los 38 g recomendados para mujeres y hombres, respectivamente, así que es posible que aún haya espacio para la mejora. De modo que si tu programa de adelgazamiento empieza a perder fuelle, quizá sea el momento de recurrir a los macrotodos y la IA.

La IA para el bien global

Una gestión efectiva y eficiente a las emergencias y a los desastres naturales, sumada a la reducción de riesgos, puede salvar vidas y viviendas, reducir al mínimo el impacto económico y aumentar la capacidad de resistencia a nivel global y local.

En los últimos veinte años, unos dos tercios de la población han sufrido desastres naturales. 4.400 millones de personas se han visto afectadas, la mayoría por inundaciones y sequías, y ha habido 1,3 millones de víctimas (muchas de ellas en terremotos), 2 billones de dólares en pérdida económicas y una gran cantidad de población desplazada, lo que ha aumentado los riesgos sanitarios, la escasez de alimento y el shock emocional tras el desastre.

En 2012, el 93 por ciento de los 905 desastres naturales registrados en todo el mundo se debieron a la meteorología, y este tipo de desastres va en aumento debido al aumento global de las temperaturas en la superficie terrestre.

Sólo el 33 por ciento de los países que sufrieron estos desastres tenían una renta per cápita baja o baja-media, pero el 81 por ciento de las víctimas pertenecían a ese grupo de países. Y la tendencia va empeorando. El número de desastres por año va creciendo (14 por ciento de aumento entre 1995 y 2004); más del doble que los diez años anteriores). Se espera que hacia 2030 vivan en los 49 países más susceptibles de sufrir desastres hasta 325 millones de personas en situación de extrema pobreza.

Hoy en día la gestión de los desastres requiere cuatro fases de actuación: mitigación, preparación, respuesta y recuperación. Durante las fases de respuesta y recuperación, la comunicación es crítica, ya que conecta a los efectivos de respuesta inmediata con los coordinadores y responsables de una multitud de organismos (que deben trabajar al unísono), así como a los afectados, a las familias y a las comunidades con los sistemas de apoyo y entre sí.

Los informes falsos y la mala información pueden agravar aún más el desastre, provocar el pánico y aumentar el alcance del impacto. En estas circunstancias son esenciales la comunicación y la información fidedigna. Si los expertos reciben una avalancha de datos mal enfocados, quedarán abrumados y no sacarán nada. Y no disponer de referencias para interpretarlos y de información a tiempo real puede dificultar la labor de los científicos informáticos que colaboran en la emergencia; es necesario disponer de los datos idóneos, en el formato idóneo, y deben llegar a las personas idóneas, con tiempo suficiente para que puedan tomar decisiones.

La convergencia de la IA, el aprendizaje automático y la tecnología avanzada de sensores remotos hace más efectiva y eficiente la gestión de desastres naturales y situaciones de emergencia. Las innovaciones en la robótica también ayudan muchísimo. Desde el 11 de septiembre se han usado robots en 47 desastres en 15 países, sean pequeños vehículos aéreos controlados a distancia, como drones, o pequeños robots tácticos terrestres capaces de colarse por donde no pueden entrar las personas ni los perros. Estos robots pueden proporcionar datos geoespaciales en grandes incendios, en misiones de búsqueda y rescate, de valoración de daños, de transporte de provisiones y medicinas, en la gestión de desastres costeros —robots marinos— y en el apoyo a los cuerpos de respuesta inmediata.

Los datos móviles también contribuyen a solucionar algunas de las necesidades más acuciantes de nuestra especie, y a mejorar el bienestar. Pero es importante contar con mecanismos transparentes para que la gente pueda compartirlos con este objetivo y retirarlos de los registros de datos agregados. Esta cuestión hace necesario contar con normas de referencia internacionales, como las que se aprueban en sectores como la banca, la asistencia sanitaria, la educación, las redes sociales y otros servicios en línea.

Necesitamos actualizar nuestros estándares técnicos, la normativa y la legislación antes de hacer uso de este nuevo tipo de datos de comportamiento humano para el bien público. Para afrontar el problema de la privacidad hay que adoptar soluciones técnicas que la protejan, y dar a los usuarios un control completo sobre la cantidad de datos que se sienten cómodos compartiendo por el bien social.

Hay numerosos ejemplos de iniciativas en que se han compartido datos, o en que las operadoras telefónicas los han recopilado de los móviles de forma anonimizada y cedido a la comunidad de investigadores para colaborar en la búsqueda de soluciones para problemas mundiales, como es el Datathon for Social Good en el Reino Unido los D4D Challenges de Orange, el Data Challenge de Telecom Italia o el acceso a los datos de Digicel por parte de los investigadores tras el terremoto de Haití en 2011.

También es esencial tener en cuenta las posibles consecuencias no deseadas que podría tener la publicación de la información obtenida del análisis de datos, aunque sea para objetivos de salud pública o para el bien social. Por ejemplo, la inferencia de patrones de migración para comprender mejor y poder prever la extensión de una

pandemia podría poner a ciertas poblaciones en riesgo en zonas de agitación social o conflicto entre diferentes grupos étnicos.

Los macrodatos y la IA pueden ser de ayuda en la lucha contra las epidemias globales y pueden aportar beneficios sociales a los más necesitados. Las pandemias globales pueden tener consecuencias catastróficas, pero con los análisis de datos móviles se podrían rebajar los índices de infección futuros y el número de víctimas mortales. En un mundo conectado en que el tráfico global de datos móviles ha crecido en nada menos que un 70 por ciento el año pasado, éstos pueden tener una aplicación que suponga beneficios para la sociedad en general.

Teniendo en cuenta los brotes de gripe aviar registrados últimamente por toda Asia, queda claro que las pandemias siguen siendo una amenaza real y persistente. Los datos móviles y la IA se pueden usar para comprender mejor los movimientos de la población durante una pandemia, para construir modelos realistas de cómo se extienden las enfermedades infecciosas y para ayudar a los gobiernos a intervenir y contener una pandemia.

La IA es clave para predecir el riesgo de una infección de transmisión humana; para calcular la difusión de la resistencia a los fármacos de patógenos como la malaria; para comprender las migraciones humanas tras un desastre natural o una situación de emergencia; y para cuantificar la exposición a agentes contaminantes del aire o a otras sustancias químicas presentes en el medio ambiente, y eso tiene una gran influencia en los programas de control y eliminación de estos riesgos de salud pública.

El enfoque tradicional aplicado en el análisis de los patrones de movilidad se basa en encuestas puerta por puerta y en la información que proporcionan los registros del censo. No obstante, estos conjuntos de datos recogidos de forma tradicional sufren de sesgos y limitaciones en lo relativo al tamaño de la muestra de población usada en el análisis, debido sobre todo al coste excesivo que tiene la adquisición de los datos. Los datos de las encuestas y de los censos proporcionan una instantánea de la dinámica de la población en un momento determinado.

En el caso de la salud pública, es de importancia capital obtener una imagen de los patrones de movilidad y de sus fluctuaciones de manera continuada, especialmente durante las emergencias (como el brote de una posible pandemia o un desastre) para poder facilitar la toma de decisiones o valorar el impacto de las medidas y restricciones gubernamentales para aumentar al máximo el impacto de cualquier intervención. En estos casos, en los centros de transporte y distribución los empleados de la sanidad pública suelen contar a la gente manualmente.

Frías-Martínez y otros científicos del banco mundial, Naciones Unidas y el MIT, han analizado esta cuestión y han empleado la IA para estudiar cómo se mueve la población. Han supuesto que los usuarios que suelen comunicarse a menudo tienen más probabilidades de estar cerca físicamente, lo que aumenta la probabilidad de transmisión de infecciones entre ellos. Su enfoque quedó validado durante el brote de

H1N1 en México en 2009, y demostró que el pico de la infección podía reducirse aproximadamente en un 10 por ciento y retrasarse aproximadamente 40 horas gracias a las acciones gubernamentales fundamentadas en la tecnología *big data* y el análisis con IA.

En 2014 nos enfrentamos al peor brote de ébola de nuestra historia. Dada la relevancia de los estudios previos y teniendo en cuenta la ubicuidad de los teléfonos móviles, un pequeño grupo de expertos en *big data* e investigadores —entre ellos un grupo de Telefónica de España— propusieron que se usaran los datos de móviles agregados y anonimizados para ayudar a luchar contra la enfermedad.

En la batalla contra el ébola, los teléfonos móviles podrían tener un valor incalculable, no sólo como herramientas para transmitir información relativa con la salud pública o para llamar a servicios de emergencias, sino también por los datos que generan.

Ya en otras ocasiones se ha demostrado la utilidad de utilizar los registros de datos para realizar un seguimiento de la difusión de enfermedades. Por ejemplo, durante la epidemia de gripe porcina mexicana, en 2009, las alertas médicas no consiguieron reducir la movilidad pese al cierre de negocios, oficinas, etc., pero los datos demostraron que las del gobierno sí lo hicieron, y eso, a su vez, redujo el número de infecciones por debajo de lo que cabría esperar.

Todo este trabajo confirma el valor que tiene evaluar escenarios hipotéticos con los macrodatos y la IA. Estas herramientas están al alcance de las instituciones públicas de todo el mundo, que pueden medir así el impacto de sus acciones en la sociedad. Y no hay motivo por el que no pueda aplicarse esta metodología a otros problemas, como las enfermedades de las cosechas, las migraciones en masa u otras amenazas de pandemia. Son amenazas que no podemos prevenir pero que, al menos, con la tecnología de hoy, podemos comprender mejor, reflejar en modelos y —quizá algún día— controlar.

Teniendo en cuenta las dimensiones del brote de ébola en África en 2014, no es de extrañar que se recurriera a los macrodatos y a la IA para combatirlo. Transformaron el campo de batalla, pasando de los cálculos aproximados y la información anecdótica a una respuesta precisa. Los recursos eran limitados, por lo que demostró ser de importancia vital en la predicción de la extensión geográfica de la enfermedad y en el envío de la correspondiente asistencia. Se empezó elaborando un mapa de móviles, permitiendo que Flowminder, organización sueca sin ánimo de lucro, mapeara los patrones de movimientos típicos de la población y determinara dónde parecía que iban apareciendo los nuevos casos. Determinando el lugar de origen de las llamadas a las líneas de asistencia a los enfermos, pudieron deducir y predecir la localización de los brotes.

Se instauraron otras medidas, como un seguimiento más específico de las personas infectadas. David Bolton, de Qlik, empresa de análisis de *big data* que también creó una *app* de seguimiento del ébola, explica que los datos de puertos,

trenes y vuelos, así como de reconocimiento de matrículas, pueden ayudar a hacer un seguimiento de las personas potencialmente infectadas y determinar con quién pueden haber estado en contacto.

Aparte del seguimiento de la evolución del ébola, el centro de control de enfermedades también necesitaba coordinar a sus efectivos. Por este motivo se dirigieron a BioMosaic, que hizo un seguimiento de la red de transporte aéreo global en tiempo real. El programa permitió sintetizar mapas agregados realizados con conjuntos de datos diversos, que incluían desde meteorológicos y climatológicos a los de distribución global de aves de corral o de los casos de enfermedad confirmados. Estos datos permitían seguir a las personas que se movían de un país a otro, o por los diferentes puertos, y le daba al centro de control de enfermedades la oportunidad de seguir a las poblaciones en mayor riesgo a través de los aeropuertos y los puestos de vigilancia, donde se hacía un control más exhaustivo. Nada de lo que se hacía en la lucha contra el ébola era tradicional. El recurso a los complejos conjuntos de datos y a su análisis es lo que permitió contener la enfermedad tan rápidamente.

En un mundo más conectado que nunca, cada vez es más fácil la dispersión de enfermedades, no sólo en un país o una región, sino también mediante rutas de viajes internacionales. El aumento de epidemias como la de la fiebre amarilla o la del Zika ponen de relieve este nuevo riesgo global. También han cambiado las expectativas sobre la rapidez a la que debería responder la comunidad internacional.

Las respuestas a las epidemias se basan en datos en tiempo real a nivel nacional y mundial. Tanto si esa información dice a los implicados en dar una respuesta donde centrar sus recursos limitados como si indica lo que piensa la población con mayor índice de riesgo sobre una amenaza determinada o la información que hay que aportar a las poblaciones afectadas, las organizaciones internacionales necesitan tener la posibilidad de saber y de actuar más rápido que nunca.

Unicef usa los datos y la IA para proporcionar una solución innovadora a los grandes retos sociales. Ahora Unicef puede asociar las necesidades de las poblaciones más vulnerables del mundo con una serie de soluciones tecnológicas y de datos y con colaboradores especializados. Estas iniciativas están en sus primeras fases, pero ya apuntan a un mundo en el que la comprensión en tiempo real de los riesgos y los desafíos globales nos permitirá trabajar mejor, y más rápido, por el bien de los niños.

Muchos hemos oído hablar de los Objetivos de Desarrollo Sostenible establecidos por Naciones Unidas para 2030 y cada vez hay más empresas y organizaciones que colaboran para que se puedan cumplir. No obstante, hay algunas empresas específicas en determinados sectores que cuentan con unos activos importantísimos que pueden ser clave para acelerar el tránsito hacia esos objetivos. Uno de esos activos son los datos y la IA.

Los datos pueden aplicarse a todos los Objetivos de Desarrollo Sostenible, usándolos para medir los progresos del sector público y privado, y para ayudar a los políticos a tomar decisiones fundamentadas y tener el mayor impacto social posible.

En una cumbre reciente para el Bien Común celebrada en Ginebra se reunieron científicos y expertos para discutir sobre la importancia de la IA en el mundo. La conclusión fue que la IA debía considerarse una herramienta que puede ayudar a la humanidad a resolver muchos problemas, con un impacto potencial en los siguientes campos:

- Eliminar la pobreza: elaborar un mapa de pobreza y establecer análisis de datos de valor predictivo.
- Cero hambre: aumentar la productividad de la agricultura.
- Salud y bienestar: analizar enormes cantidades de datos de asistencia sanitaria.
- Educación de calidad: revolucionar las aulas con el aprendizaje personalizado.
- Igualdad de género: detectar las desigualdades e impulsar el empleo equilibrado entre ambos sexos.
- Agua limpia y eliminación de residuos: mejorar y hacer más eficiente la provisión de agua limpia.
- Energía limpia y asequible: mejorar la producción de energía fotovoltaica.
- Trabajo decente y crecimiento económico: aumentar la productividad con la automatización inteligente.
- Innovación e infraestructuras en la industria: contribuir a impulsar la innovación en la industria.
- Reducir la desigualdad: crear una sociedad más inclusiva (por ejemplo, con la robótica para discapacitados).
- Ciudades y comunidades sostenibles: una mejor planificación urbana usando los datos de los sensores para tomar decisiones.
- Consumo y producción responsables: predicción de los niveles de producción óptimos para reducir los residuos.
- Actuación ante el clima: crear modelos de cambio climático para predecir los desastres.
- Vida submarina: controlar la pesca ilegal con software de reconocimiento de patrones.
- Vida en la tierra: ser más listos que los furtivos y hacer un seguimiento de la salud de las especies.
- Paz, justicia e instituciones fuertes: reducir la discriminación y la corrupción en el gobierno.
- Asociación para obtener objetivos: la colaboración multisectorial es esencial.

La IA puede ayudar a los emprendedores y a los gobiernos del siglo XXI a construir cambios sociales progresivos pero expeditivos, por ejemplo impulsando el uso de los teléfonos móviles con IA aumentada para la monitorización de la salud, tal

como ha propuesto Margaret Chan, directora general de la Organización Mundial de la Salud.

La IA ha despertado un interés en el poder de los datos y el análisis como motor de la innovación, el aprendizaje y la búsqueda de nuevas perspectivas. Ha hecho que nos demos cuenta del uso que tienen los datos y las herramientas de análisis para comprender problemas complejos. Ha introducido el análisis cuantitativo en el razonamiento. Ha abierto grandes horizontes al descubrimiento. Los macrodatos no sólo tienen un papel que jugar en la aceleración del aprendizaje y en las soluciones de negocio: si se aplican de forma responsable, los *big data* y la IA pueden usarse para afrontar todo un abanico de complejos desafíos globales en aras del bien común.

IA ética, privacidad y el alma de los datos

La gente está perdiendo el gusto por la vida privada y ya ni la percibe.

MILAN KUNDERA

La utopía del ciberespacio

A mediados de los años ochenta surgió en el mundo una nueva idea. La idea de que un nuevo sistema, un sistema invisible, podía gestionar la sociedad a través de un mundo de ordenadores con gigantescas redes de información. Esas redes de información dieron origen a nuevas estructuras y sacudieron el status quo existente. Fue el inicio del ciberespacio, un espacio en el que un grupo de visionarios situaban el sueño de una nueva utopía, de un mundo en el que podría reinventarse el viejo mundo.

Los utópicos tecnológicos de la Costa Este de Estados Unidos concibieron el ciberespacio como un mundo nuevo y seguro en el que podrían hacerse realidad una serie de sueños radicales. Frente a la complejidad de la política, los utópicos de la informática vieron en el ciberespacio un lugar en el que podían refugiarse de la política y de las dificultades del día a día y construir un mundo más justo de equilibrio global. Aquellas redes de máquinas e información empezaron a crear una nueva realidad alternativa que vivía en una dimensión paralela al mundo real. El ciberespacio podía liberar a cualquiera de las viejas estructuras políticas y de poder, y permitía soñar con un futuro nuevo, brillante y más sencillo.

Uno de los exponentes de esta idea fue John Perry Barlow. Barlow había escrito mucho explicando las maravillas de internet. Para él internet era más que una red de ordenadores. Era lo que llamaba una «frontera electrónica». Él veía internet como un lugar en el que podías imaginarte que descubrías un continente tan vasto que quizá no tuviera fin. Un lugar donde podías imaginarte un nuevo mundo con más recursos de los que puede llegar a agotar nuestra codicia en el futuro, más oportunidades de las que pueden llegar a aprovechar los emprendedores, y un sector inmobiliario peculiar que se expande a medida que evoluciona.

Organizó cibermaratones para intentar dar forma al ciberespacio. También escribió un manifiesto en defensa de la independencia del ciberespacio, dirigiéndose a los políticos para que no pusieran las manos en él. Pensó en internet como un lugar

libre y mágico, una poderosa imagen de una red libre del control de las corporaciones, y alternativa a las viejas estructuras de poder.

Y esa visión dominó internet durante más de dos décadas. Internet sería un lugar donde cualquiera podría expresar sus convicciones, un lugar independiente y libre por naturaleza.

Pero dos jóvenes hackers de Nueva York destrozaron la visión de Barlow. Eran personajes de culto en la red, y se hacían llamar Phiber Optik. Lanzaron ataques contra las redes de ordenadores gigantes que consideraban que ocupaban un lugar central en la nueva estructura informática e intentaron demostrar que ese nuevo mundo era un mundo de fantasía.

Pensaban que esta nueva tecnología de comunicación global presentaba graves defectos a la hora de dar «el poder a la gente». Para una nueva generación que crecía en un mundo menos jerárquico y más interconectado, la posibilidad de producir y compartir su información, en un espacio de acceso libre, resultaba algo natural. Pero los hackers no estaban de acuerdo.

En una célebre discusión en línea, los dos hackers atacaron a Barlow. Les enfurecía que Barlow estuviera convencido de que en el nuevo mundo del ciberespacio no habría jerarquías ni poderes que lo controlaran, algo que ellos estaban convencidos de que no sería verdad. Así que decidieron demostrar que Barlow se equivocaba.

Los dos hackers se colaron en un gran sistema informático de una sociedad llamada TRW (ahora Experian) que había sido usado por el ejército estadounidense durante la Guerra Fría. En aquel momento TRW se usaba para guardar los registros financieros de millones de estadounidenses de modo que pudieran usarse para tomar decisiones sobre préstamos y créditos. Los hackers se hicieron con el historial crediticio de Barlow y lo publicaron en la red.

Con ello los hackers demostraban el poder de estos nuevos bancos de datos, y que podían usarse para desafiar al idealismo del ciberespacio a través de un enorme conjunto de sistemas de información *big data*. Los hackers planteaban así la posibilidad de que la utopía del ciberespacio de Barlow pudiera ocultar un tipo de poder diferente basado en los datos y la inteligencia para procesarlos que, a menos que se mantuviera equilibrado, podría tener consecuencias inesperadas. Tanto Barlow como los hackers tenían razón. Para llevar a término la visión de una internet abierta, en la que lo principal fuera la gente y no unos nuevos poderes centralizados, había que conseguir un equilibrio sostenible.

La privacidad de los datos

A pesar de todo el progreso tecnológico, a veces a mí también me entran ganas de apearme durante un tiempo de internet, y veola fragilidad de la visión inicial de

internet que tenía Barlow como una red donde todo es bueno. Sí, me he pasado toda mi vida profesional trabajando en tecnologías de internet y buscando el modo de hacer que la distribución de datos sea más escalable y más segura. Pero me levanto por la mañana y leo twitter, Facebook, blogs, periódicos, el correo electrónico, agregadores de noticias RSS, recibo diferentes tipos de publicidad y recomendaciones y empiezo a tener una sensación de urgencia, de predestinación, de direccionamiento, el peso de las expectativas sobre lo que debería hacer a continuación; siento el sesgo sobre el tipo de información que consumo o el modo en que debería pensar. A veces tengo la sensación de que he externalizado parte de mi pensamiento crítico, cediéndoselo a la nube, a los macrodatos y a la IA y que, con ello, he perdido parte de mi libertad.

Hoy en día una gran parte de internet se ha convertido en una enorme máquina de rastreo de datos personales. Hace tanto tiempo que es así que raramente le damos relevancia, pero vale la pena recordarlo de vez en cuando. Toda esta libertad de comunicación y de información supone que prácticamente cada sitio que visites te colocará una docena de cookies de rastreo, lo que permite hacerte un seguimiento de un sitio a otro con contenidos personalizados. La personalización requiere de datos, así que esas cookies están por todas partes. Los sitios web intentan que no te agobies con ellas, algunos más que otros, pero no hay quien escape a la tendencia dominante.

Lo que es aún peor es que si esa información y datos personales caen en manos de cibercriminales, el impacto personal y a la sociedad puede ser muy importante. Espionaje, chantajes, o abusos en Estados totalitarios puede ser uno de los riesgos de que los personales, que hoy se guardan en centros de datos seguros, caigan en manos inadecuadas. El estado de Corea del Sur admite tener más de 1.700 ciberhackers para conseguir información personal de usuarios que hoy confiamos en que está bien protegida. Esto puede empeorar cuando más y más cosas estén conectadas a internet. Por ejemplo infraestructuras públicas críticas pueden convertirse en armas de ataques masivos a la población si caen en las manos equivocadas. Internet permite acceso remoto a las cosas e infraestructuras, y cada vez es más importante que los datos que generan las cosas conectadas a Internet estén protegidos. Desafortunadamente, algunos de estos ataques están pasando más y más a menudo. Ahora pasa a pequeña escala, pero un día puede pasar a gran escala. El origen de la economía de los datos, y por tanto del riesgo de que caigan en manos equivocadas está en el comercio electrónico y la personalización.

Si coges el teléfono, abres un navegador y buscas en la sección de vídeos de Google «España», ¿qué ocurre? Bueno, en la mayoría de casos, lo más probable es que aparezca un pequeño videoclip antes de empezar a ver el vídeo principal: un anuncio publicitario. Y es muy posible que cada persona vea uno diferente, aunque todos estén esperando que se cargue el mismo vídeo.

¿Pero cómo se ha conectado el teléfono móvil con un vídeo? Primero el teléfono conecta con un *router* inalámbrico local o una estación base de móviles. Ese *router*

está conectado con la red de acceso local. Esa red está conectada con la red nacional, que a su vez está conectada con una de las principales redes internacionales.

Y desde allí el teléfono se conecta con el vídeo, que está alojado en el centro de datos. Muchos de estos centros se encuentran en Islandia o en países nórdicos; eso sirve para rebajar el coste de la refrigeración, necesaria 24 horas al día, 7 días por semana, 365 días al año. Prácticamente son unas neveras enormes y muy caras que analizan tus datos utilizando la IA y el aprendizaje automático para ofrecerte sugerencias o recomendaciones y personalizar tus contenidos.

Internet está llena de anunciantes y proveedores de contenidos. No obstante, a diferencia de los medios tradicionales como la televisión, los periódicos, la radio o las revistas, la publicidad y los contenidos en internet están muy personalizados, basándose en tus datos personales. ¿Qué significa eso? Los medios tradicionales se basan en estadísticas agregadas para realizar marketing de productos, el concepto de que pocas compañías de gran envergadura (empresas de venta a gran escala, de deportes, de ocio) quieren llegar a un gran público con un anuncio interesante, una publicidad «de pocos para muchos».

Este marketing *offline* tradicional se basa en datos agregados de grandes tendencias de consumo y estadísticas anónimas, y no requiere conocer tus hábitos personales exactos, por ejemplo qué periódicos has leído antes del que estás leyendo en ese momento, o qué emisoras de radio has sintonizado antes de poner la tele.

En internet el juego es otro. La publicidad y el marketing en internet se basa en datos muy personalizados y pueden acceder a ellos muchísimas empresas, no ya únicamente compañías como Coca Cola o Nike, sino también el tipo que vende tablas de surf customizadas por internet, o las agencias de viajes que proponen vacaciones en cualquier lugar del mundo, sobre todo pequeñas empresas que intentan llegar a los consumidores interesados en algún campo o producto que cubra sus necesidades personales. De hecho, los diez principales anunciantes de todo el mundo proporcionan a Google menos del 5 por ciento de sus ingresos por publicidad. Gran parte de los anuncios vienen de pequeñas empresas que se dan a conocer en internet.

Pero para asociar contenidos de marketing y clientes, se necesita conocer las necesidades y las intenciones de la gente, sus datos personales. Si no, las grandes empresas de publicidad se perderían en un mar de comerciantes y usuarios incapaces de entrar en contacto unos con otros. Y para hacer esta asociación, para comprender las necesidades y las intenciones de la gente, las empresas de publicidad recurren a cosas como los hábitos de navegación, los patrones de movilidad, las búsquedas en internet, etc. Al alma de tus datos.

Personalizar un contenido o una publicidad según cada usuario no es el objetivo final. El objetivo final es que compres más por internet, en sitios de comercio electrónico como Amazon, e-Bay, etc., donde una vez más, según tus datos, obtendrás recomendaciones de productos de electrónica, películas, ropa o cosas que han gustado a tus amigos o que han comprado.

A estas alturas muchas personas han comprendido que vivimos en un nuevo tipo de economía de datos. Los «revendedores de información», también conocidos como *data brokers*, han recopilado cientos o miles de detalles —lo que compramos, nuestra etnia, nuestras preocupaciones económicas y de salud, nuestras actividades en la web y en las redes sociales— prácticamente de todos los adultos del mundo. Otras empresas especializadas en hacer *rankings* de consumidores usan algoritmos de IA para hacer valoraciones de los usuarios de internet, identificando a algunos como consumidores «valiosos», a los que vale la pena ofrecer tarjetas de crédito en mejores condiciones u otros beneficios.

Hay otro tipo de compañía, llamada plataforma de *ad-trading*, que traza un perfil de los usuarios de internet y concede acceso a unos o a otros subastándolo, en una práctica conocida como «pujas en tiempo real». Es algo que se ha sabido hace poco, y varios miembros del Congreso de Estados Unidos y alguna Agencia Federal ya han iniciado investigaciones sobre el asunto o alrededor del debate sobre Cambridge Analytica.

La huella digital que van dejando los consumidores en diversos canales y medios crece exponencialmente, y por fin se dan cuenta de que hay quien comercia con esos datos para obtener un beneficio económico, creando potenciales sesgos de información y condicionándonos; en vez de ayudar a la gente a encontrar un nuevo espacio de libertad tal como imaginó Barlow. Esto, combinado con la creciente preocupación sobre la seguridad de los datos y los ciberataques, hace que la gente quiera saber, cada vez más, cuándo se recogen registros suyos, qué es lo que se almacena y quién lo hace, y cómo se usan después esos datos.

Por una parte, puede que no nos sintamos nada cómodos sabiendo que usan el alma de nuestros datos para presentarnos determinados contenidos en internet. Pero hay que tener en cuenta que estas mismas técnicas te ayudan a encontrar lo que quieres y necesitas sin tener que perder horas y más horas buscando por la web.

Pero cuando estás en internet, a veces da la sensación de que hay un desajuste fundamental entre la economía de la atención, y las aspiraciones que tiene cada uno. A veces da la impresión de que la economía de la atención nos está transformando el cerebro. Con su flujo incesante de ofertas, de oportunidades para las relaciones, sus interrupciones y sus interacciones, la economía de la atención cambia nuestros límites físicos y psicológicos y convierte los impulsos en hábitos, a veces nocivos.

La mayoría de anuncios incentivan los diseños que buscan más sacar partido a nuestra atención que a nuestras intenciones. (Podría debatirse si un anuncio que respete y apoye la intención del usuario debería considerarse realmente «publicidad»). Y como las interfaces digitales son mucho más maleables (por el simple hecho de que se basan en software) de lo que han sido nunca la televisión o la radio, los entornos digitales se pueden adaptar mucho mejor a la lógica del diseño de la publicidad. Cuando no existía el software, la publicidad siempre era la excepción a la regla, pero ahora, en el mundo digital, la publicidad se ha convertido en la regla.

Pero te daré una pista: probablemente no vaya dirigida a hacer posible ninguno de los objetivos que te has planteado tú mismo. Tus objetivos quizá sean «pasar más tiempo con los niños», «aprender a tocar la guitarra», «perder diez kilos antes del verano», «licenciarme», etc. El tiempo de que dispones es limitado, y lo sabes.

Estas tecnologías, por otra parte, intentan maximizar objetivos como «tiempo pasado en el sitio», «número de visualizaciones del vídeo», «número de visitas a la página», etc. De ahí los cebos, de ahí los vídeos que se reproducen solos, de ahí la avalancha de notificaciones. Tu tiempo es limitado, y estas tecnologías lo saben.

De hecho, esto ha provocado que los bloqueadores de publicidad se hayan convertido en uno de los servicios más populares entre los usuarios de ordenadores personales y dispositivos móviles. Los medios generalistas también están aumentando la cobertura de este tema en sus primeras páginas y en los horarios de máxima audiencia, en ocasiones haciendo asombrosas revelaciones sobre el uso de los datos personales, y los órganos reguladores de ambos lados del Atlántico empiezan a movilizarse para afrontar los desafíos relacionados con la privacidad.

En cierto sentido, los datos, la IA e internet están diseñándote la vida, y eso debería preocuparte. Lo que nos parece diversión (videoclips, mensajes, servicios de datos, etc.) puede llegar a alejarnos de nuestras propias necesidades, nuestros propios objetivos, nuestro objetivo vital, la capacidad de reflexionar y liberarnos de lo que ya no nos resulta útil, para fijar la atención y las energías en las cosas que nos sirven para algo.

Internet y la tecnología podrían ayudarte a pensar con mayor eficiencia, pero al mismo tiempo también pueden provocar que tu cerebro funcione de un modo menos eficiente creando adicciones, haciéndote más dependiente y condicionándote, apartándote de la visión de Barlow. Nuestro cerebro puede recibir numerosas interrupciones en forma de pequeñas informaciones, SMS, correos electrónicos o publicidad que atasquen la mente y la hagan menos efectiva. Para el cerebro eso supone una ráfaga constante de interrupciones breves que nos impiden hacer razonamientos a largo plazo, y darnos cuenta de que necesitamos ser nosotros mismos y pensar libremente. Básicamente, nuestro cerebro está perdiendo el tiempo en pensamientos inútiles y superficiales, tendencias y modas, y todo ello a costa de nuestra privacidad.

Cuando aún no había internet, si alguien te hacía una pregunta complicada, tenías unas cuantas opciones: podías ver si alguien que conocieras sabía la respuesta. Podías consultar una enciclopedia. O podías dirigirte a la biblioteca y hacer una investigación. Cualquiera que fuera la opción escogida, sin duda sería más complicada y te llevaría más tiempo que la que aplicarías hoy en día: buscar en Google.

Gracias a la tecnología —y a internet en particular— ya no dependemos de nuestros recuerdos, a veces poco fiables, para rescatar hechos puntuales e informaciones particulares. Piénsalo: ¿Cuándo fue la última vez que te molestaste en

memorizar el número de teléfono de alguien? ¿Y qué sentido tiene aprender cómo se escribe esa palabra tan larga y complicada si el autocorrector se encargará de que la escribas bien?

Pero ahora que tenemos todo el conocimiento que podamos llegar a necesitar en la punta de los dedos, ¿estamos externalizando nuestra memoria, confiándosela a internet? Según algunos estudios recientes, sí. El último, realizado por investigadores de las universidades de California e Illinois, concluyó que al confiar cada vez más en internet está cambiando nuestro modo de pensar y de recordar.

En el estudio se pidió a dos grupos de personas que respondieran a una serie de preguntas de cultura general. A los del primer grupo se les pidió que usaran solamente la memoria, mientras que los otros tenían que buscar las respuestas en internet. Luego se hizo una serie de preguntas sencillas a ambos grupos y se les dio la opción de usar internet. Entre los que habían usado internet la primera vez fueron muchos más los que volvieron a hacerlo.

No sólo consultaron internet en una mayor proporción, sino que también lo hicieron más rápidamente, sin intentar apenas deducir la respuesta por sí mismos, aun cuando las preguntas eran relativamente simples.

Todo esto pone de manifiesto una tendencia que los investigadores denominan «descarga cognitiva». Se ha vuelto tan fácil buscar cualquier cosa en internet que estamos incluso dejando de recordar o de procesar determinadas cosas. Antes quizá intentáramos recordar algo por nosotros mismos; ahora ni nos molestamos. A medida que disponemos de cada vez más información vía *smartphone* y otros dispositivos, vamos confiando cada vez más en ellos para cualquier cosa de nuestra vida diaria.

Este último estudio se suma a investigaciones anteriores que sugieren que internet no sólo está cambiando nuestro modo de vivir y de trabajar: está incluso alterando nuestros cerebros. Nicholas Carr, autor de *¿Qué está haciendo Internet con nuestras mentes?*, cree que confiando en internet como disco duro externo de nuestra memoria, estamos perdiendo la capacidad de transferir los hechos que oímos y leemos a diario de nuestra memoria operativa a nuestra memoria a largo plazo, algo que según Carr es «esencial para la creación de conocimiento y sabiduría».

Los datos y la tecnología también nos cambian por dentro. Perdemos cada vez más la paciencia con la realidad, especialmente cuando es aburrida o incómoda. Acabamos esperando cada vez más del mundo, y lo queremos más rápido. Y como la realidad no puede satisfacer nuestras expectativas, aumenta la frecuencia con que deseamos volver a la pantalla, y en muchos casos tomamos decisiones erróneas. Es un bucle automantenido por retroalimentación.

Vamos a necesitar desarrollar cada vez más nuestros propios criterios para decidir con qué quedarnos, qué leer y qué pasar por alto en internet, dejando así espacio para nuestros propios pensamientos y razonamientos.

No obstante, cada vez cedemos más alegremente toda nuestra información, todos nuestros datos; compartimos lo que comemos, lo que hacemos, lo que somos. Esto

tiene el riesgo de que acabemos en una situación en la que hackers usen todos estos datos de manera nociva y maliciosa, o simplemente privándonos de la capacidad de tomarnos el tiempo necesario para reflexionar y tomar nuestras propias decisiones. Esto, llevado al extremo, sería una sociedad como la que predijo Huxley en *Un mundo feliz*, en la que la gente alcanza la felicidad a costa de no disponer de ninguna intimidad ni control sobre su vida, sacrificando la privacidad a cambio de placeres a corto plazo.

Yo, como individuo, creo que podemos reivindicar el derecho de disponer de una vida privada y decidir cuándo y por qué no hacerla pública, comprendiendo el precio y el valor de la privacidad.

Prácticamente todas las cosas a las que nos apuntamos requieren nombres, números y direcciones de correo electrónico, y a menos que nos leamos atentamente los términos y condiciones, no tenemos ni idea de dónde acabarán esos datos.

Y no es de extrañar, dada la extensión y la complejidad de los términos y condiciones de privacidad que imponen la mayoría de empresas. Recientemente el grupo de presión australiano Choice hizo una demostración práctica del problema, haciendo que un actor leyera las 73.198 palabras de los términos y condiciones de Amazon Kindle. Tardó nueve horas. Según Stephanie Hankey, cofundadora y directora ejecutiva del Tactical Technology Collective —cuyo Data Detox Kit ayuda a los consumidores preocupados a limpiar su huella digital—, es uno de los problemas típicos a los que se enfrentan los usuarios de tecnología, que se encuentran una y otra vez con documentos de términos y condiciones «ininteligibles e inaccesibles», aunque eso no es más que la punta del iceberg.

«A diferencia de los datos para la prestación de servicios [como la localización para Google Maps], hay ocasiones en que la recogida de contenidos y metadatos forma parte del modelo de negocio de una empresa, que luego los vende —advierte Hankey—. Se han dado bastantes casos de extralimitaciones. Por ejemplo, la *app* de la linterna que se apropia de los contactos de tu agenda. Está claro que no los necesita para darte luz».

Cada vez digitalizamos más todos los aspectos de nuestras vidas, y en muchos casos dejamos una huella increíblemente detallada en la red. ¿Cómo nos aseguramos de que nuestra información esté segura? ¿En quién podemos confiar?

Nueve de cada diez personas no tiene ni idea de qué hacen las empresas con la información personal que tienen de sus clientes. En un estudio reciente realizado por el Chartered Institute of Marketing (CIM) con 2.500 personas, resultó que el 57 por ciento no confía en que ciertas empresas traten sus datos con responsabilidad. Y el 51 por ciento se quejaba de que habían sido contactados por empresas que habían usado sus datos de forma no deseada.

El CIM afirma que las políticas de datos personales de los sitios web deberían ser más claras y más simples. Dice que las empresas no consiguen convencer a la gente

para que lea los términos y condiciones de modo que comprendan lo que se hace con sus datos personales.

Sin embargo, de acuerdo con su investigación, la gente no tiene ni idea de lo que hacen con sus datos, y tiene miedo de los hackers, que les timen o les inundan de spam. El CIM encuestó a más de 2.500 consumidores y profesionales del marketing para ayudar a las empresas que buscan nuevos modos de vender nuestros datos por internet.

Pero en lugar de descubrir a un ejército de consumidores bien informados contentos de compartir sus datos, el estudio reveló un muro de perplejidad y desconfianza por parte del público.

Más del 92 por ciento de los que respondieron no entendían del todo cómo se usaba la información que obtenían las empresas de ellos, y se mostraban muy escépticos con respecto a las prácticas de marketing.

Por ejemplo, el 71 por ciento de los consumidores no se sentía cómodo con que hubiera servicios que siguieran sus movimientos a través de los *smartphones*. Sin embargo, el número de aplicaciones que recopilan estos datos de geolocalización son ya el 20 por ciento. A la mayoría de personas no les gusta compartir informaciones de sus perfiles de las redes sociales. Y sin embargo, el 44 por ciento de las empresas los recopilan. Lo peligroso es que la mayoría de los que navegan por internet aceptan los términos de un sitio web sin enterarse de cómo se va a usar cualquier información que compartan.

Tim Berners Lee, el inventor de la web, definió recientemente las tres cosas que había que cambiar para salvar la web. Y la primera es que tenemos que recuperar el control de nuestros datos personales, porque en su opinión lo hemos perdido.

Berners Lee afirma que el modelo de negocio actual de muchos de los sitios web consiste en ofrecer contenidos gratuitos a cambio de datos personales. Muchos de nosotros damos nuestro consentimiento: aunque en muchos casos sea aceptando largos y confusos documentos de términos y condiciones, fundamentalmente no nos importa que recojan cierta información nuestra a cambio de servicios gratuitos. Lo malo es que estamos pasando algo por alto.

Nuestros datos van a parar a almacenes de datos, donde son procesados por IA, lejos de nuestro control, y con ello perdemos las ventajas de que podríamos disfrutar si tuviéramos un control directo sobre ellos y pudiéramos decidir cuándo y con quién compartirlos. Es más, a menudo no tenemos ningún modo de indicar a esas empresas qué datos preferiríamos no compartir —especialmente con terceros—: los términos y condiciones son «o todo o nada».

Y si un hacker se hace con esos datos, el daño puede ser mayor. De hecho, esta recolección de datos tan extendida puede afectar a la privacidad de otros modos. En regímenes represivos es fácil ver el daño que puede causar: la detención de blogueros y el seguimiento de los oponentes políticos. Pero incluso en países con gobiernos más avanzados, el que todo el mundo esté controlado puede ser un jarro de agua fría para

la libertad de expresión, y puede impedir que se use la web para debatir asuntos importantes, como los temas sensibles relacionados con la salud, la política, la sexualidad o la religión. Hoy en día el control de los datos y lo que vemos en internet puede afectar elecciones y la propia democracia, como hemos visto en las recientes incursiones de hackers en la escena política en Estados Unidos y Europa.

Pero la misma tensión que está pasando ahora con los datos y su libre circulación por internet —sin mucho control por el usuario— ya pasó hace unos años, de otra manera, con los contenidos de vídeo/música y las redes Peer-to-Peer (P2P). Y a mí me tocó vivirlo en primera persona...

Un delicado equilibrio

Es una mañana fría en Redmond, la ciudad al sur de Seattle donde tiene su sede central Microsoft. Es el invierno del año 2002 y hace sólo unos meses que me acabo de incorporar a los prestigiosos laboratorios de redes y sistema distribuidos de Microsoft en Cambridge. Estoy en Seattle para verme con los científicos liderando el diseño de la red de Microsoft para distribuir contenidos a todos los ordenadores que tienen instalado el sistema Windows. Llego a la oficina conduciendo un coche automático de gran tamaño y difícil de maniobrar, hundido en el asiento después de haber intentado ajustar su altura sin mucha suerte, y aparco en el parking de visitantes. Una de las primeras personas que voy a ver pertenece al equipo directo de Bill Gates, para contarles mi visión de cómo las redes P2P pueden ayudar a acelerar que el contenido viaje de forma rápida y libre por internet y escuchar como esto puede ser útil para Microsoft y sus usuarios de Windows.

Antes de entrar al edificio 34, y mientras espero, empiezo a leer una vez más el manifiesto escrito por John Perry Barlow declarando la independencia del ciberespacio sobre el mundo actual. Según estoy llegando al final del texto, me quedo mirando esta frase: «En nuestro nuevo mundo del ciberespacio, sea lo que sea lo que la mente humana pueda crear, puede ser reproducido y distribuido infinitamente sin ningún coste». Y empiezo a darme cuenta que el concepto que Barlow está describiendo, la capacidad de que una idea, un contenido, o un dato, pueda ser distribuido de manera casi instantánea e infinitamente por la red, está en la base de las redes P2P en las que yo estoy trabajando! Se llaman P2P porque en el intercambio de archivos y contenidos no interviene un tercero, es de un cliente a otro cliente y por tanto la libertad de copiar contenidos de manera rápida por la red es mucho mayor.

Al no existir una autoridad única que pueda colapsarse, la red tiene una capacidad innata de rapidez y sobrevivir por sí misma, además de tratarse de una conexión fuerte y muy versátil. A pesar de tener grandes ventajas, las redes P2P también tienen

su parte mala. Ese consumo de ancho de banda puede llegar a ser excesivo, ralentizando otros procedimientos que se hagan con el ordenador que está actuando. Además, se pueden utilizar para violar los derechos de propiedad intelectual, haciendo que la industria creativa pierda el control de sus contenidos, que circulan de manera rápida y sin límites por la red.

Después de mis reuniones en Redmond, entiendo mucho mejor que la visión de Barlow de poder liberar las ideas y los contenidos por internet, y las redes P2P que yo estoy diseñando, pueden tener grandes beneficios y aplicaciones. Pero al mismo tiempo, pueden encontrar oposición en las industrias existentes de la información y el contenido que pierden el control de sus contenidos y datos. De hecho entiendo que muchos artistas, creadores y compañías productoras de contenidos están preocupadas por el crecimiento de las redes P2P de intercambios de archivos que de manera rápida cogen el contenido y lo distribuyen por internet sin fronteras ni límites. Son redes donde el trasvase global de información ya no necesita de otras formas físicas de distribución (CDs, DVDs, cintas de vídeo); pero no todo el mundo está de acuerdo con esta visión de un mundo donde la distribución de los contenidos no tiene ningún tipo de control.

De vuelta en Cambridge, y cuatro años después, con mi bicicleta y de camino a casa después de haber acabado el trabajo en el laboratorio, recibo una llamada desde las oficinas centrales de Microsoft en Redmond. El sistema P2P que hemos creado (Avalanche) y que Microsoft ha adoptado para distribuir antivirus, contenidos y actualizaciones a gran velocidad por internet, ha creado grandes oportunidades para parar ataques informáticos y mantener los ordenadores de todo el mundo más al día. Pero también está creando gran preocupación entre la industria cinematográfica y de la música. Esta industria ve un riesgo en que este mismo sistema P2P súper-rápido de Microsoft que hemos creado se utilice de forma equivocada por hackers o piratas informáticos y los artistas y productoras acaben perdiendo más aún el control de sus contenidos (música y películas) de manera más extendida y rápida.

Después de recibir la llamada descubro que la BBC, CNN, y otros grandes medios están hablando de la posibilidad de que nuestra red P2P sea usada para acelerar la piratería de contenidos, cosa para la que no ha sido diseñada. Y esto está causando gran revuelo en la red, los medios de comunicación y en la industria en general.

Llego a mi casa, una casa adosada de estilo Victoriano cerca del pub Egale, el pub donde Francis Crick y James Watson se reunían para descubrir los secretos de la vida y empezar la decodificación del ADN. Pero esta vez, me quedo poco tiempo porque tengo que preparar las maletas para pasar los próximos meses en la Costa Oeste de Estados Unidos, entendiéndome qué cambios tenemos que hacer para que esta red P2P

que he desarrollado para distribuir contenidos, incluya también las preocupaciones de los propietarios de los contenidos para permitirles el control y garantías que necesitan, usando más métodos de encriptación y gestión de derechos de datos.

Y es ahí, en ese momento donde me doy cuenta yo mismo que el sueño de Barlow de un ciberespacio independiente y libre para las ideas, los datos y los contenidos es eso, una utopía. Una utopía que necesita tener a todos en cuenta, que tiene dos caras, y que debe respetar el control de la información y los datos por parte de sus creadores.

Y hoy, quince años después, me doy cuenta de que ese mismo reto que pasó con las redes P2P y la libertad de compartición de los contenidos por internet vs. su control, está pasando hoy con los datos, los algoritmos de IA y su control. Donde cada vez más personas se preocupan por saber dónde están sus datos, qué se hace con ellos, y cómo protegerlos. Y no tanto cómo pueden viajar de manera libre por internet, expuestos al riesgo del mercadeo de datos, ciberataques, hackers, o fugas y filtrados de información.

Edward Snowden es un experto en informática estadounidense que antes trabajaba para la NSA, la Agencia de Seguridad Nacional de Estados Unidos. La NSA recopila montones de datos y los procesa para reducir el riesgo de sufrir potenciales ataques a la seguridad nacional. Todos los países tienen un equivalente a esta agencia. Pero en 2013, y sin autorización, Edward Snowden actuó como un hacker, copiando, fugando y revelando información clasificada de la NSA. Sus revelaciones pusieron en evidencia numerosos programas de vigilancia global, muchos gestionados por la NSA en colaboración con grandes corporaciones.

Los documentos filtrados revelaron la existencia de un proyecto secreto llamado PRISM, y demostraron que PRISM era un programa inteligente de recolección de datos con el que la NSA podía tener acceso a datos personales en empresas privadas. A la opinión pública no le hizo ninguna gracia.

Así que alguien tenía que contar la otra versión de la historia. Esta responsabilidad recayó en el general Keith Alexander, que hasta hace poco ha sido director de la NSA. Su misión era demostrar que había un elemento de compensación. Si tienes la privacidad en un lado y la seguridad en el otro, y quieres más seguridad nacional, tienes que comprometer la privacidad.

Y ahí resultó que la NSA era una máquina eficiente que convertía los datos personales en seguridad nacional. El general Alexander dijo que esos programas de recopilación de datos eran vitales para frenar las amenazas terroristas. Y era innegable que éstos en poder de empresas como Facebook eran cruciales para dirigir las iniciativas de seguridad nacional.

La privacidad, algo de lo que antes sólo se preocupaban las elites, se está convirtiendo en un tema de Estado por un lado y de conversación común por el otro. Hace unos años los consumidores no se preocupaban de su privacidad o de sus datos personales, o no tenían los conocimientos necesarios al respecto, pero los estudios

recientes demuestran un gran aumento entre los que querrían saber más sobre las opciones disponibles de privacidad y de control de datos personales. La gente se preocuparía más de su privacidad si supieran hasta qué punto están expuestos en la red.

Al interactuar en línea y navegar por internet vamos dejando rastros de datos. Dejamos rastros de datos propios, con nuestros números y nuestras acciones, nuestra posición, nuestros gustos y aversiones, del mismo modo que los insectos dejan rastros. Igual que los insectos, que dejan rastros de feromonas, en tiempos de crisis estas señales nos pueden llevar a otras personas que comparten nuestras preocupaciones y que tienen la voluntad de ayudar, por ejemplo, revisando registros de conexión, el correo electrónico, información sobre nuestra actividad en las redes sociales o nuestras compras.

Y el problema es que, como seres humanos, tenemos limitaciones. Tenemos un cerebro limitado. Tenemos una capacidad limitada. A veces no podemos tomar siquiera las decisiones más simples sobre nosotros mismos. Tomamos decisiones basándonos en razonamientos erróneos, y muchas veces no tenemos tiempo siquiera para pensar o razonar. Tenemos que recurrir a los presentimientos, a la intuición, y seguir adelante. Los datos y la IA podrían servirnos de ayuda.

Los datos y la IA podrían convertirse en tu nueva guía, en tu nueva intuición. ¿A qué ritmo crees que puedes generar esos rastros de datos? Como individuos, recibimos datos de nuestros sensores físicos: los ojos, los oídos, la nariz, el tacto. Todos esos sentidos obtienen un input que pasa por nuestro sistema nervioso central y se codifica. Algunos de esos eventos se procesan, otros se almacenan, y otros se descartan.

Hace unos diez años, Gordon Bell, científico que trabajaba en la zona de la bahía de San Francisco, decidió registrar y almacenar todo lo que hacía en forma digital. Decidió hacer una instantánea de todo que sucediera en su vida que tuviera interés usando una cámara siempre encendida. Lleva encima un podómetro, un pulsómetro y un grabador de voz. Tiene un registro de todos sus emails, de sus vídeos; sabe todo lo que pasa a su alrededor.

Uno podría preguntarse: ¿Cuánto ocupan esos datos? ¿Cuánto son diez años de vida de Gordon Bell? Bueno, esos diez años caben en un lápiz de memoria USB de 64 GB que se puede llevar en el bolsillo y que cuesta menos de 20 euros. ¡Una década de tu vida cabe en un pendrive!

Sí, hay muchos datos. Pero ¿dónde están todas esas feromonas de datos? Antes estaban en tu PC. Ahora, cada vez que haces una transacción por internet, con una red social, con tu correo electrónico, cuando compras algo, toda esa información está ahí y revela quién eres y lo que haces. Y está en la nube, las nuevas refinerías de datos, que del mismo modo en que extraemos petróleo y lo transformamos en productos químicos, plásticos, fármacos, queroseno, toman éstos y los transforman con la IA y

el aprendizaje automático en consejos y recomendaciones de salud, de estilo de vida, de fitness y de desarrollo personal.

A pesar de los adelantos algorítmicos en la anonimización de los datos, se ha demostrado que es factible deducir identidades a partir de datos de comportamiento humano, en particular cuando se combinan con otros de diferentes fuentes. Por ejemplo, el equipo del profesor Zang ha demostrado que si se dispusiera de las direcciones de casa y del trabajo de algunos usuarios, se podría des-identificar hasta un 35 por ciento de los usuarios de la red simplemente usando sus localizaciones más frecuentes (que probablemente serían la de casa y la del trabajo). Montjoye y su equipo han llevado esta idea más lejos y han demostrado que los datos de movilidad GPS de cada individuo son únicos, y que pueden usarse para desidentificar a los usuarios con una precisión del 95 por ciento.

Es necesario desarrollar protocolos de datos compartidos anonimizados similares a los que usa desde hace tiempo la comunidad médica. Además, la comunidad científica ha propuesto diversas tecnologías para potenciar la privacidad, como la ofuscación, que consiste en alterar ligeramente la información de los datos de un modo irreversible para que no refleje la localización real, pero que siga siendo representativa del fenómeno estudiado; y el k-anonimato, que asegura que las trayectorias individuales sólo puedan revelarse si hay al menos k-1 trayectorias idénticas a la trayectoria específica que se pretende compartir y analizar.

Otro campo explorado actualmente por empresas como Apple consiste en determinar el número máximo de niveles que se pueden exponer de un individuo, una agregación de datos espaciales y temporales, que permita hacer modelos precisos e inferencias preservando todo lo posible su privacidad: encontrar patrones de datos sin examinar estos últimos. No obstante, aún nos faltan años para conseguir diseñar algoritmos de IA que puedan funcionar con datos encriptados, aunque sin duda es el futuro al que debemos apuntar.

El rápido avance comercial en las capacidades de IA plantea una serie de cuestiones importantes, desafiando a la sociedad para garantizar la no discriminación y comprender la toma de decisiones en los sistemas automatizados. Los responsables políticos, y los reguladores están preocupados por la posibilidad de que IA tenga un sesgo a la hora de tomar decisiones automatizadas. De hecho, cada vez se oye más hablar sobre la equidad algorítmica.

Las herramientas actuales de inteligencia artificial basadas en el aprendizaje automático pueden incluir los prejuicios de sus creadores y programadores, o el de las personas que se usan como conjuntos de datos de entrenamiento y prueba de los algoritmos. En la siguiente etapa de IA, cuando le demos la capacidad de tomar decisiones importantes en nuestra vida (por ejemplo, si nos debemos operar o no), entonces, se vuelve aún más crítico poder auditar y ajustar de forma transparente las cuestiones éticas de tales sistemas de decisión, desarrollando la ética digital y empatía digital.

Por último, hay una creciente alarma de que la complejidad del aprendizaje automático puede reducir la justificación de las decisiones consiguientes de «el algoritmo me obligó a hacerlo». Muchos algoritmos de IA emplean tal complejidad que ni siquiera sus diseñadores pueden explicar o interpretar cómo llegan a las respuestas. Las decisiones se toman de una «caja negra» y deben tomarse por fe. Eso puede no importar si IA recomienda una película, pero hay más en juego si IA maneja un automóvil, ofrece un diagnóstico de salud o recomienda un próximo trabajo.

Ahora mismo no parece que a la gente le preocupe tanto la privacidad; sin duda no tanto como para sacrificar alguna de las ventajas prácticas que ofrece la web. Pero todo puede cambiar. Muchas otras cosas no suscitaban preocupación en el pasado: el fumar en lugares públicos, la seguridad en carretera, la seguridad en los aeropuertos, los juguetes peligrosos o la discriminación racial o sexual, por ejemplo. Las sociedades evolucionan.

Por mucho éxito que tengan la ciencia de los datos y la IA, no son indestructibles. También pueden caer en desgracia. La pérdida de confianza es el talón de Aquiles de la IA y del uso de *big data*.

Prácticamente la mitad de los usuarios de internet afirma que la preocupación por la privacidad y la seguridad no les ha frenado a la hora de hacer cosas básicas en internet —como colgar posts en las redes sociales, expresar opiniones en foros o incluso comprar cosas en sitios web—, según una reciente encuesta realizada por el gobierno de Estados Unidos y publicada en 2016.

Aun así, tal como muestra un estudio realizado con 41.000 unidades familiares que usan internet en Estados Unidos, el efecto disuasorio que crea este tipo de preocupación demuestra que la inseguridad de los datos empieza a tener consecuencias que van más allá de los efectos secundarios de que el individuo pierda sus datos personales. La investigación sugiere que algunos consumidores están llegando a un punto en que sienten que ya no pueden confiar en internet para sus actividades diarias.

Cada día, miles de millones de personas de todo el mundo usan internet para compartir ideas, realizar transacciones económicas y mantenerse en contacto con la familia, los amigos y los colegas. Pero para que internet crezca y prospere, los usuarios deben seguir confiando en que su información personal se mantendrá segura y en que se protegerá su privacidad.

El estudio demostró que casi el 20 por ciento de los sujetos había experimentado personalmente algún tipo de robo de identidad, una violación de seguridad en línea u algún otro problema similar durante el año anterior a la realización del estudio. Y el 45 por ciento del total afirmó que sus preocupaciones por la privacidad y la seguridad en la red les había hecho dejar de usar la web con finalidades prácticas.

Cuando se les pidió que enumeraran sus mayores motivos de preocupación, prácticamente dos de cada tres participantes citaron el robo de identidad, y casi la

mitad el fraude con tarjetas de crédito o datos bancarios. Uno de cada cinco mencionaron la recogida de datos privados por parte del gobierno.

Otros estudios han demostrado que cada vez se hacen más comunes las preocupaciones sobre la privacidad. En un estudio realizado por Pew Research se observó que la gran mayoría de personas quiere controlar sus datos digitales, pero tiene poca confianza en poder llegar a proteger esa información.

El estudio también demostró que un número significativo de estadounidenses han intentado proteger su privacidad en la red de diferentes modos. Casi el 60 por ciento afirmaba haber eliminado las cookies o el historial de su navegador, y prácticamente uno de cada tres decía haber borrado o editado algo posteado previamente en las redes.

Pero muy pocos manifestaron que se hubieran planteado usar métodos más avanzados para proteger su privacidad, como la encriptación del correo electrónico. Y otro estudio de Pew Research ha demostrado que a la mayoría de usuarios les cuesta mucho decidir cuáles son las mejores herramientas y estrategias para proteger su privacidad significativamente.

Los nuevos datos de la Agencia Nacional de Telecomunicaciones de Estados Unidos sugieren que un número considerable de estadounidenses se han decantado al menos por una estrategia: evitar las actividades en la red. Esa tendencia podría tener graves consecuencias para los bancos, el comercio electrónico y la economía de la red en general. Además de ser una cuestión que preocupa a muchos estadounidenses, los problemas de privacidad y seguridad pueden llegar a reducir la actividad económica y dificultar el libre intercambio de ideas en línea.

El problema es que podemos perder el control de la privacidad y ser objeto de discriminación en razón de los datos. Las investigaciones han demostrado que podemos ver diferentes internet según el alma de nuestros datos y los algoritmos de IA, una para los ricos, una para los pobres, una para los que usan Mac, otra para los que tienen un PC, si los billetes de vuelos baratos quedan escondidos en la segunda página de resultados para los usuarios de Mac o si los mismos zapatos aparecen un 20 por ciento más caros porque se da el caso de que acabas de echar un vistazo a la revista Forbes.

De hecho, ser anónimo en la red es más difícil de lo que puede parecer, tal como ha revelado un reciente estudio publicado en *Frontiers in ICT*. Tus hábitos de navegación pueden indicar tu personalidad y generar una firma digital única que te identifique, a veces con sólo navegar media hora.

«Nuestra investigación sugiere que se pueden deducir los rasgos de personalidad de una persona a partir del uso que hace de internet en general. Esto difiere de otros estudios que sólo han analizado el uso de plataformas sociales como Facebook o Twitter», explica el doctor Ikusan R. Adeyemi, investigador de la Universiti Teknologi de Malasia y director de este nuevo estudio.

Tradicionalmente, las investigaciones que relacionan rasgos de personalidad con el uso que se hace del ordenador suelen centrarse en los medios sociales. Por ejemplo, la gente extrovertida tiende a usar estas plataformas para aumentar su esfera de amistades e influencias, mientras que los individuos introvertidos pasan más tiempo en las redes sociales para compensar una probable falta de interacción física.

No obstante, los hábitos de navegación de una persona y la IA también pueden reflejar sus opciones, sus preferencias y sus reflejos, controlados en gran medida por sus características psicológicas particulares.

Este último estudio reclutó a voluntarios de la Universiti Teknologi de Malasia y monitorizó el uso que hacían de internet, incluyendo muchos factores, como la duración de las sesiones de internet, el número de sitios web visitados y el número total de consultas realizadas.

Por otra parte, los voluntarios respondieron a un test para determinar sus características de personalidad en cinco categorías: apertura a nuevas experiencias, meticulosidad, extroversión, complacencia e inestabilidad emocional. El análisis de estos datos reveló una intensa relación entre la personalidad de la persona y su conducta de navegación. Por ejemplo, con sólo treinta minutos de navegación podía distinguirse el nivel de conciencia del individuo.

«Los datos pueden usarse para ajustar productos a un público específico; pueden usarse para desarrollar un servicio de internet inteligente capaz de predecir y personalizar la experiencia del usuario —explica el doctor Adeyemi—. También pueden usarse como método complementario para aumentar la seguridad en la identificación, la autenticación en la red, y problemas de la delincuencia en la red».

El doctor Adeyemi y sus colegas de la Universti Teknologi esperan que sus investigaciones futuras permitan determinar nuevos patrones de los cinco rasgos de personalidad investigados con el fin de poder trazar un mapa de personalidad del individuo y aumentar así nuestra comprensión de los hábitos de navegación en internet.

El mal uso de los datos y la IA podrían acabar agotando el recurso más valioso —la confianza del público—. Esto podría llevarnos al ya conocido efecto de la tragedia de los comunes, con la retirada de los consumidores al tener la sensación de que su privacidad, sus datos y el control de sus decisiones corren riesgo.

Así pues, nos encontramos en un punto delicado, en el que pueden coincidir todos los elementos de la tormenta perfecta, amenazando la privacidad de la gente. En los años venideros necesitaremos proteger el alma de nuestros datos y la transparencia y la ética de IA con un nuevo paradigma que nos asegure un mayor control sobre nuestros datos y sobre el uso que se les da.

Ya hay quien está tomando medidas para que así sea. Por ejemplo, el Media Lab del MIT está creando «sistemas de seguridad» para los datos personales. Y luego tenemos el Data Transparency Lab, una ONG en la que científicos, defensores de una internet abierta e instituciones públicas de todo el mundo desarrollan herramientas,

servicios y nubes de datos personales para ayudarnos a recuperar el control sobre el alma de nuestros datos.

Según la Wikipedia, la Tragedia de los Comunes es una teoría económica ideada por Garret Hardin que afirma que los individuos que actúan de forma independiente y racional siguiendo su propio interés actúan en contra del interés de la colectividad, agotando recursos comunes. El término deriva del título de un artículo escrito por Hardin en 1968, que a su vez se basa en un ensayo de un economista victoriano sobre los efectos del pastoreo no regulado en territorio comunitario.

Nikos Laoutaris, director científico del Data Transparency Lab, hace una analogía entre la Tragedia de los Comunes y el tira y afloja que presenciamos actualmente en la web entre privacidad y datos e IA.

En la Tragedia de los Comunes original, cada uno de los ganaderos actuaba de un modo egoísta, dejando pastar cada vez más vacas en un terreno comunitario, aun sabiendo que un número desproporcionado de animales acabaría dejando el terreno sin hierba y que inevitablemente supondría el fin del negocio de todos ellos. Todos los granjeros compartían ese conocimiento, y sin embargo, no hacían nada para evitar el evidente desastre que les esperaba.

Para explicar esta paradoja, hay que tener en cuenta el egoísmo y la capacidad de autoengaño del ser humano. El egoísmo dicta que es mejor obtener beneficios inmediatos y aumentar el número de reses, dirigiendo los daños provocados a los demás y posponiendo las consecuencias, dejándolas para el futuro. El autoengaño hace referencia a la creencia utópica de que se pueden ir acumulando cada vez más vacas sin llegar a afrontar la tragedia porque, milagrosamente, los demás se moderarán y reducirán el tamaño de sus respectivos rebaños, salvando así el campo de la destrucción. Desgraciadamente, todo el mundo piensa igual, por lo que las reses acaban con toda la hierba del campo, destruyéndolo.

La analogía que se puede establecer entre la metáfora ficticia de las vacas y la preocupación general, por la erosión de la privacidad y el potencial abuso de la IA es muy real.

¿Qué podemos hacer para evitar una «tragedia de los comunes» en el campo de IA y la privacidad? «El mejor desinfectante es la luz del sol». Quizá haya llegado el momento de aplicar esta famosa cita de Louis Brandeis, fiscal del Tribunal Supremo de Estados Unidos, a los desafíos que plantea la defensa de los datos y la ética de la IA. Esto pasaría por proveer de más transparencia de cómo los algoritmos de IA toman decisiones y usan tus datos.

Hace unos años, un grupo de científicos de la Universitat Politècnica de Catalunya demostró que la discriminación de precios ya parece haber encontrado hueco en los algoritmos de IA del comercio electrónico. Eso significa que el precio que una persona ve en su navegador al buscar un producto o servicio puede ser diferente al que ve otro usuario en ese mismo momento en una localización diferente.

Incluso estando en el mismo sitio, los rasgos personales de un usuario —entre ellos su historial de navegación— pueden provocar variaciones en el precio ofrecido.

En ese mismo sentido, un grupo de investigadores de la Universidad de Columbia, la Northeastern University y el INRIA han desarrollado herramientas y metodologías con los que los usuarios finales pueden comprobar si los anuncios y recomendaciones que han recibido están dirigidos a ellos específicamente, si tienen una variable aleatoria o si dependen de la localización.

Este tipo de herramientas mejoran la transparencia del uso que se hace en la red de los datos personales, lo cual tiene múltiples beneficios para todas las partes.

En el pasado, la transparencia ha demostrado ser bastante efectiva como medio para llevar internet hacia la dirección correcta. El desarrollo de herramientas de transparencia para la privacidad y la protección de los datos sin duda contribuirá a que la Tragedia de los Comunes siga siendo una metáfora y no una realidad en el ecosistema de IA.

Bancos de datos personales

Bueno, ¿y qué hacemos ahora? Internet hoy en día tienen una habilidad para recopilar datos mucho mayor incluso que la de los servicios de inteligencia. Por ejemplo, el servicio de reconocimiento de rostros de Facebook que se usa para etiquetar las fotos de los usuarios puede ser útil para cualquier espía que quiera determinar quién es una persona determinada y quiénes son sus amigos, partiendo de una simple foto.

Así pues, ¿deberíamos educar a nuestros hijos para que usaran únicamente aplicaciones creadas en países que consideremos amigos, o simplemente deberíamos pensar que vivimos en un mundo que sencillamente no tiene secretos? Sea como sea, una cosa está clara: no hay nada que sea gratis. Y en internet tampoco. O pagamos con dinero, o con nuestros datos y nuestra privacidad.

El combustible más valioso del mundo ya no es el petróleo, sino los datos. Y una economía de datos e IA exige un nuevo enfoque con respecto al tratamiento que damos a éstos.

Los gobiernos podrían potenciar la aparición de nuevos servicios abriendo más sus propios almacenes de datos, o gestionar partes cruciales de la economía de éstos del mismo modo que gestionan las infraestructuras públicas, tal como hacen la India con su sistema de identidad digital o Estonia con los registros de datos de sus ciudadanos. También podrían exigir que ciertos datos se compartieran —con el consentimiento de los usuarios—, enfoque que se está imponiendo en Europa, donde a los bancos se les exige que hagan accesibles a terceros los datos de los clientes.

Lo importante es que cualquier persona que quiera procesar la totalidad del alma de sus datos probablemente necesite que esté toda en un mismo lugar de su propiedad y sobre el que tenga control. Y puede que la evolución del sistema actual nos lleve a

la aparición de nubes privadas para los personales de cada uno, bancos de datos personales donde podrías tener juntos todos tus datos, almacenarlos, replicarlos, hacer *backups* y copias de seguridad.

En muchos casos el mecanismo para utilizar tus datos personales es un complejo «Acuerdo de Licencia con el Usuario Final», un documento de más de 32 páginas que permite acceder a tus datos personales. Por supuesto, cualquiera puede optar por usar un servicio diferente si tiene objeciones. Pero no hay muchas alternativas.

La alternativa más definitiva es optar por abandonar el mundo de la red, pero eso resulta sencillamente imposible para la mayoría de las personas. El modo de plantearse la competencia en el pasado, en la época del petróleo, queda desfasado en lo que ha dado en llamarse la «economía de los datos». Necesitamos un nuevo enfoque.

En su libro *¿Quién controla el futuro?*, Jaron Lanier, científico informático y uno de los pioneros de la realidad virtual, plantea que convenciendo a los usuarios para que proporcionen información valiosa sobre sí mismos a cambio de servicios «gratuitos», hemos vendido nuestra alma digital. La solución que él propone es hacer que las transacciones en línea sean bidireccionales (y que se compense a los usuarios) para asegurarnos de que los individuos se den cuenta del valor económico que tienen sus datos personales, que ahora mismo ceden sin más.

Por su parte, Doc Searls, uno de los estadistas más veteranos de la vieja internet, que actualmente trabaja en el Berkman Centre de Harvard, hace un planteamiento similar en su libro *The Intention Economy: When Customers Take Charge*, pero propone una solución diferente. La idea básica es que del mismo modo que la mayoría de grandes empresas usan los sistemas de «gestión de relaciones con los clientes» para gestionar sus interacciones con los usuarios, los clientes necesitan nuevos sistemas que puedan gestionar sus interacciones con las empresas, pero defendiendo los intereses del cliente.

La filosofía de base de todos estos intentos por nivelar el campo de juego es la creencia de que los datos del individuo le pertenecen a él, y que nadie debería tener acceso a ellos a menos que sea de forma controlada por el propietario de los datos. Por tanto sigue en marcha la búsqueda de tecnologías (software y/o hardware) que lo hagan posible y que resulten fáciles de usar.

Llegar a contar con un servicio que pueda usar cualquier individuo sin duda llevará tiempo y esfuerzo. Entre los problemas que hay que resolver está el espinoso asunto de la confianza, la seguridad y la facilidad de uso. Pero se puede llegar muy lejos si damos el primer paso, siguiendo el ejemplo de Lanier y Searls.

Empresas de muy diversa índole podrían responder ofreciendo servicios que ofrezcan un mayor control a los consumidores. Entre ellas podrían contarse compañías como Yahoo, Microsoft y Google, que ya almacenan determinadas categorías de datos y los hacen disponibles para sus clientes; o redes sociales como Facebook y LinkedIn.

Podrían buscarse otras soluciones en los bancos de datos personales como el que proporciona la *startup* Personal, que permiten que los consumidores almacenen y gestionen determinados tipos de datos personales como si estuvieran en un almacén de datos personal; o empresas como <Reputation.com>, dirigida a clientes dispuestos a pagar por gozar de privacidad de éstos.

Usando estos bancos de datos, los consumidores podrían elegir entre diversas empresas o instituciones para compartir o almacenar diferentes categorías de información personal. Podrían seleccionar una institución financiera para que protegiera los económicos, un centro médico para que gestionara los de salud, y un servicio de seguridad de datos de consumo como gestor que los aglutinara todos. Luego, cuando una persona estuviera dispuesta a comprarse un coche, podría autorizar a su almacén de datos personal para compartir los económicos o de seguros relevantes con la empresa vendedora. O podría autorizar a su aseguradora para que revisara los de consumo y aumentara automáticamente la cobertura de su seguro de hogar cuando se comprara artículos caros, como un sistema de *home cinema*.

Los bancos de datos podrían dar un mayor control a los consumidores sobre quién ve determinados tipos de información y cómo se usa. En Estonia ya se está instaurando un sistema de este tipo, que custodia una gran cantidad de datos de los ciudadanos, que aun así pueden compartirse con determinadas partes (por ejemplo, policía, gobierno, sistema de salud) con permiso del usuario, e informándole cada vez que alguien accede a algún dato crítico.

Estonia ya ha demostrado una gran habilidad en la creación de servicios gubernamentales en línea. Es pionera en el uso de identidades digitales protegidas para sus ciudadanos, lo que les permite firmar y encriptar documentos, acceder a servicios gubernamentales y realizar compras en línea.

Las bases de datos públicas y privadas de Estonia se comparten a través de una red de pares llamada X-Road, una especie de federación de la información. Los usuarios dan su consentimiento digital usando su carné de identidad y su PIN, para que una base de datos extraiga información de otra (por ejemplo, si un hospital necesita contactar con un seguro médico para saber el estado de cobertura de un paciente). Las personas que crearon el sistema descentralizado estonio afirman que el encendido debate sobre privacidad de los datos que se libra en Estados Unidos está muy mal enfocado. Habría que centrarlo en dar a la gente el control necesario para que decida quién tiene acceso a sus datos.

En el futuro es posible que tengas todo tu historial de datos personales —el alma de tus datos— en un lugar, un banco de datos, que te ayudará a sacar partido a éstos y te ofrecerá servicios útiles. Y del mismo modo que instalamos *apps* en nuestros teléfonos móviles, podremos descargar *apps* en nuestros bancos de datos personales, aplicaciones que usarán nuestros datos para utilizarlos en nuestro beneficio, aplicaciones que funcionarán con el alma de nuestros datos, con una representación digital de nosotros mismos.

Estos bancos de datos personales almacenarán información sobre hábitos de navegación, registros de localización, conductas de consumo, detalles económicos como informes financieros, contactos de correo y de redes sociales, así como apuntes en el calendario, etc. Y serás tú, como propietario de éstos, quien decidirás cómo y cuándo compartirlos con diferentes empresas y terceros, para que te proporcionen nuevos servicios basados en el alma de tus datos personales.

De hecho, los datos personales —el alma de tus datos— son mucho más valiosos que los agregados. Cuanto más privados/íntimos y más cercanos al contexto del usuario, más valiosos se vuelven. La diferencia de valor entre un perfil anónimo de datos y un perfil personal muy detallado, con hábitos y preferencias de consumo, puede ser de cientos de dólares, o puede multiplicarse por 500.

Es muy importante que conservemos el control de esos datos personales, que comprendamos dónde se encuentran, quién tiene acceso, por qué motivo, y cómo los usan. Tenemos que controlarlos para gestionar mejor nuestras vidas. Tenemos que ser capaces de denegar el acceso a esos datos, darlo cuando nos convenga, borrarlos si nos conviene y tener un nivel de confianza y de transparencia que nos dé la confianza necesaria para seguir compartiendo nuevos datos en línea.

A pesar de que los datos y la IA pueden proporcionarnos grandes oportunidades para potenciar el bien común, la salud pública y mejorar la vida de la gente, aprovecharlos no es en absoluto algo trivial. El almacenamiento, el acceso y el procesamiento de datos con información personal sensible, como la localización y la movilidad, las conexiones a internet, los patrones de mensajería, así como la información relacionada con la red social del individuo, deben seguir las leyes de protección de datos, y un claro código ético de conducta y transparencia.

Ya existen diversos proyectos en esta dirección. Por ejemplo, el gobierno del Reino Unido ha lanzado una gran iniciativa para devolver los registros de cliente estructurados a los individuos a cada interacción que hacen los usuarios con las empresas. Hablamos de los datos de las facturas de suministros, o de los bancarios.

Por su parte, empresas como la operadora O2 han empezado a dar acceso a los usuarios a sus datos personales para que puedan controlarlos, por ejemplo a su historial de localización, de búsquedas o de llamadas telefónicas. Y por último están las iniciativas para crear bases de datos personales, como el Data Locker, que es un archivo privado de código abierto. Puedes echar un vistazo al código, ver cómo se mantienen seguros los datos, cómo se comparten, cómo puedes controlarlos, y compartirlos si quieres con tus amigos mediante una red de pares.

Telefónica, también planea devolver a sus clientes de telefonía los datos que ha recogido para que puedan usarlos y conservarlos ellos mismos —aunque cambien de operador— con la esperanza de que dando un nuevo enfoque a un asunto tan delicado como la privacidad y el control de éstos potencie la fidelidad del cliente.

Probablemente algún día tengamos que empezar a pensar en escribir un testamento de datos para nuestro banco de datos personal, del mismo modo que lo

hacemos para disponer de nuestras propiedades físicas. Los datos que custodiarán esos bancos serán tan importantes que tendremos que pensar bien qué nos quedamos, qué borramos, qué queremos compartir y con quién.

Por ejemplo, puedes compartirlos con tu esposa para que sepa lo que sientes en un momento dado. O compartirlos con tus hijos para que entiendan qué es lo que te fue bien y lo que no; tus éxitos y tus fracasos. Y —¿por qué no?— podrías donar tus datos a la ciencia, del mismo modo que donas tus órganos hoy en día, para crear un enorme corpus de datos de la condición humana que puedan estudiar las próximas generaciones.

Los datos pueden constituir un enorme laboratorio de la condición humana. Y del mismo modo que donamos nuestros órganos para que los médicos los estudien de modo que puedan encontrar la manera de curar enfermedades, quizá en el futuro los científicos estudien tus datos privados para alcanzar nuevos avances científicos.

En 2017 Estonia abrió la primera «embajada de datos» del mundo en Luxemburgo, un almacén para guardar una copia de seguridad de todos los datos de Estonia, que gozará de los mismos derechos soberanos que una embajada cualquiera, pero que podría servir para «reiniciar» el país a distancia en caso de un ataque masivo de piratería informática.

La respuesta de Tim Berners Lee al desafío que se nos plantea es la de disociar los datos de los sitios web, incluidas las redes sociales. Para ello dirige el proyecto Solid del MIT, que propone «una serie de convenciones y herramientas para crear aplicaciones sociales descentralizadas basadas en los principios de datos vinculados».

Otra empresa que busca reequilibrar la situación es Datacoup, sitio web que ofrece la posibilidad de vender los datos personales propios y sacar dinero con ello. Matt Hogan, cofundador y consejero delegado, afirma que ya han atraído a «muchos usuarios que han ganado miles de dólares vendiendo sus datos».

Para Martijn Verbree, socio de la empresa de seguridad informática KPMG, las iniciativas para devolver la propiedad de los datos a los consumidores son «nobles pero arduas». Una vía sería imponer normas y en particular replantearse el Reglamento General sobre Protección de Datos (GDPR) de la Unión Europea.

Actualmente la oficina nacional de protección de datos del Reino Unido está realizando consultas sobre el formato y la puesta en marcha de la regulación, que debería aplicarse a partir de mayo de 2018. «La GDPR aplicará mano dura —afirma Verbree—, con multas de hasta el 4 por ciento de los ingresos del grupo o 20 millones de euros». Con la actual normativa de protección de datos, la multa más alta que se puede imponer es de 500.000 libras esterlinas.

De momento, las leyes son la mayor arma que tenemos contra el uso fraudulento de los datos personales. Pero lo más efectivo en la lucha contra la pérdida de control de los propios es sin duda la educación y contar con una mayor transparencia.

Yo estoy convencido de que podríamos estar metiéndonos en una tormenta perfecta. Una tormenta perfecta es una situación en que una combinación de

circunstancias agravan una situación drásticamente. En meteorología, un ejemplo podría ser un sistema de bajas presiones que se mueve en dirección a un sistema de altas presiones que avanza en sentido contrario. Una tormenta de estas características podría hundir hasta el barco más grande.

Es importante tener un control mayor sobre los propios datos, y devolverlos a los usuarios, de modo que mientras vivan puedan sacarle un beneficio, o disfrutar de servicios que les ayuden a conocerse mejor y crecer. Y al morir deberías poder ceder tus datos personales a tus seres queridos. O podrías ceder el alma de tus datos a la ciencia para que los investigadores los usaran, junto a otros, para buscar soluciones a los problemas del mundo.

Muy pronto los datos y la IA podrían ayudarnos a modelar políticas comunitarias, productos financieros, la asistencia sanitaria, tomar decisiones de vida o muerte por nosotros, y la educación de acuerdo con nuestras necesidades personales. Pero ¿estamos dispuestos a sacrificar nuestra privacidad a cambio de productos y servicios que nos beneficien?

Cuando la gente usa la web lo que hace es algo realmente íntimo y personal. Recurren a su médico en busca de una segunda opinión, pero consultan la web en busca de una primera, para entender qué es lo que les sucede. Comunican de un modo muy íntimo con sus familiares. Pero a veces la gente hace cosas en la web y con la IA, que dicen cosas de ellos que ellos mismos desconocen.

Dado que hoy en día gran parte de lo que hacemos en nuestra vida diaria pasa por internet, la red puede ser increíblemente reveladora. Del mismo modo que tenemos derecho a ir a un médico y tener una conversación privada cara a cara, deberíamos poder navegar por la web y usar IA y tener una experiencia igual de privada y ética.

Yo creo que en los próximos años nos preguntaremos: ¿Qué datos usa la IA y para qué? La experiencia se va volviendo más personalizada cada vez. Con los macrodatos y la IA damos relevancia al contexto pero ¿hasta dónde podemos llegar?

No tenemos que perder de vista la dimensión descomunal que van adquiriendo los datos. En los próximos dos meses reuniremos más información que toda la existente anteriormente. El aprendizaje automático y la IA hacen que la nube sepa cada vez más de ti y de lo que es importante para ti.

Esta abundancia de datos cambia la naturaleza de la competencia. Los gigantes de la tecnología siempre se han beneficiado de los efectos de la red: cuantos más usuarios tiene una empresa, más atractivo resulta para los demás hacerse usuario. Con los datos se crean nuevos efectos en la red. Al recopilarlos, una empresa tiene más posibilidades de mejorar sus productos, con lo que atraerá a más usuarios, generará más datos, etc.

Cuantos más datos reúna Tesla con sus coches sin conductor, más efectiva será la IA que hace que conduzcan solos. Eso explica en parte que la empresa, que no vendió más que 25.000 coches en los primeros tres meses de operación, tenga ahora un valor

superior al de General Motors, que vendió 2,3 millones de unidades. Un archivo de datos enorme puede actuar como foso de protección.

En unos años, nuestra vida diaria será completamente diferente. El trabajo, la educación, la propia definición de lo que es estar presente serán diferentes. Pero la gente sigue sin dar importancia a su privacidad, la ética de los algoritmos de la IA, o a sus datos privados. La gente quiere saber que nadie traspasa ninguna línea roja, pero no tiene ni idea de dónde están esas líneas rojas. Si toda esa información se usara y se gestionara correctamente, podríamos crear un mundo mejor.

Te pasas los días y las noches en las redes sociales y en la web. La imagen que nos presentaba Barlow era la de un nuevo paraíso abierto, en el que se compartiría información de forma gratuita. Pero ahora los algoritmos tienen tanto poder, y saben tanto de ti, que sólo te dan lo que saben que te gustará. Podrías quedarte atrapado en una cámara de resonancia, en la que lo único que ves y que oyes es a ti mismo. Con cada clic aumenta la fuerza de los datos de la web y de la IA. Pensamos que estamos expresándonos libremente, pero puede que acabemos convirtiéndonos en meros componentes en una burbuja de información.

Puede que se esté acercando una tormenta. Pero eso no significa que el barco vaya a hundirse. Tenemos que prepararnos ahora, planteando un debate abierto sobre qué tipo de IA queremos. Tenemos que entender a fondo las repercusiones de la apertura de lo que pueden parecernos unas oportunidades maravillosas, pero que en realidad chocan con la visión original que tenía Barlow del ciber-espacio como un lugar independiente y libre. Y tenemos que dar un paso atrás, parar un momento y ocuparnos del alma de nuestros datos, y de la ética de IA y los algoritmos, para vivir en un mundo en el que los consumidores tengan un mayor control sobre sus vidas.

Deberíamos dar a nuestros datos y a la IA el trato que le damos a nuestro espíritu: escuchar qué es lo que nos cuentan, y quererlos como nos queremos a nosotros mismos, porque pueden ser lo que ponga al descubierto nuestra propia alma, y así ayudarnos a crecer.

Conclusión

La ciencia y la ciencia ficción se influyen mutuamente. A mí, ambas me han influenciado personalmente a largo de mi vida. Siendo niño y a medida que iba creciendo, intentaba siempre crear, construir o retocar cosas, montándolas de manera ligeramente distinta para ver qué pasaría o desmontándolas para seguir aprendiendo. Crecí en La Felguera, Asturias, en una sociedad en medio del cambio. Las décadas anteriores a mi nacimiento habían traído consigo avances tecnológicos monumentales —el viaje del hombre al espacio, los satélites, los aviones supersónicos, así como la aparición de la informática y los lenguajes de programación—. Y la década en la que nací estaba a punto de traer aún más avances, como los ordenadores personales o la telefonía móvil.

La popularización de series de televisión como *Star Trek* y películas como *El planeta de los simios* crearon un nuevo tipo de superproducción cinematográfica. Estimulados tanto por la ciencia como por la ciencia ficción de nuestro tiempo, una generación de científicos e ingenieros crecimos haciéndonos preguntas como «¿y si?» y «qué es lo siguiente?».

Hoy estamos viviendo gran parte de lo que soñábamos cuando éramos niños. El espacio ya no es la última frontera e incluso esas «comunicaciones» que vimos en las películas y series de ciencia ficción son parte de nuestra vida cotidiana. Y la prueba de esto es que probablemente tengas en tu bolsillo ahora mismo un teléfono con la misma capacidad computacional que el ordenador que llevó al hombre a la Luna.

Como tecnólogos somos capaces de hacer predicciones bastante acertadas sobre los avances en los próximos cinco años; sin embargo, mirar a diez o quince años vista requiere la imaginación y el espíritu aventurero de un niño, combinados con la mente de un científico que conoce los límites de las leyes fundamentales del Universo. Perseguir lo imposible es dejar que la imaginación te lleve a lugares a los que nunca pensaste que podrías llegar, pues saber que nuestra única limitación es nuestra imaginación nos da la sensación necesaria de optimismo para poder hacer cualquier cosa.

Así que me imagino en el 2045. Me dirijo al aparcamiento y me doy cuenta de que con el auge de los coches sin conductor, hace tiempo que las autoescuelas han desaparecido. Sacarse el carnet de conducir se ha vuelto obsoleto y más un *hobby* para románticos que una necesidad real. Las autoescuelas se han convertido en centros de formación que, mediante cursillos de un día, enseñan a sacar el máximo provecho a las nuevas funcionalidades de los vehículos inteligentes (por ejemplo, entretenimiento, planificación de ruta y vacaciones, teleconferencias, etc.).

Mi coche autónomo se está acercando. El coche ha leído mi deseo de ir a casa a través de un pequeño implante de comunicación cerebral que tengo bajo la piel del

cráneo. También llevo varios sensores incrustados debajo de la piel para el control de la salud, como la presión arterial, balance de bacterias, o los niveles de oxígeno.

Entro en el coche y la interacción con éste es casi humana, en cuanto a voz, emoción y experiencia. El coche se ha aprendido mis preferencias musicales, de temperatura e iluminación, que se ajustan sin necesidad de apretar un botón.

La tecnología de control por pensamiento es aún bastante reciente, pero en mi vehículo ya está en marcha. Es la misma tecnología que utilizan los veteranos de guerra para mover sus extremidades protésicas de forma remota. Es un mundo en el que, con el mero pensamiento, puedes controlar tu entorno, de la misma manera que la neurotecnología ayuda a un paciente paralizado a comunicarse con el mundo que le rodea mediante la actividad neuronal.

Y cuando llego a casa, estos mismos avances neurotecnológicos me permiten «hablar» con la arquitectura electrónica de mi casa sin decir una palabra. Los mandos a distancia desaparecieron hace tiempo y los aparatos caseros también funcionan a través del pensamiento.

Al mismo tiempo el mundo de la nanotecnología está transformando nuestras vidas de manera impensable. Piel artificial, células solares pulverizables, partes del cuerpo autoreparables, capas de invisibilidad, y gran cantidad de aplicaciones médicas a nivel de ADN que podrán crear y reconstruir partes del ser humano y el medio ambiente. Por ejemplo las impresoras en 3D ya imprimen ADN sintético para repararnos desde el interior y regenerar partes del cuerpo como un diente autosustitutivo.

Abro la nevera y encuentro unas fresas que todavía están ácidas. Mi médico, que se ha especializado en *big data* y análisis de la información recogida por mis implantes y sensores, me ha indicado que tengo que estabilizar el nivel de PH, bajar mi colesterol, y cuidar las intolerancias. Para esto no me ha recetado un medicamento sino un conjunto de *apps* que se han convertido en los nuevos medicamentos preventivos.

Al lado del microondas tengo una nueva máquina que imprime comida en 3D con cientos de cartuchos de diferentes ingredientes y que uso para imprimir comida de manera rápida con la cantidad exacta de calorías y nutrientes que necesito.

Me acerco al piano, coloco mis guantes táctiles que están llenos de sensores y actuadores y que son capaces de generar impulsos de corriente que mueven los músculos de mis dedos, y me preparo para una clase de música. Mi profesor es el avatar de un famoso pianista fallecido, que consigue impartirme una clase magistral haciendo que mis dedos se muevan para aprender a tocar el piano de forma adecuada.

Antes de ir a dormir, tengo que planear mi próximo viaje. Me siento delante del ordenador y empiezo a controlar a un personaje de videojuego que está caminando por un mercado. Excepto que no es un juego. Es un robot que camina por las calles de Tokio, con señal de vídeo y audio. Es un robot de alquiler que me ayuda a comprobar el siguiente destino antes de aterrizar y desde el anonimato. Los robots se han

convertido en el áter ego de las personas, entendiendo sus emociones y psicología, ayudándoles a ejecutar tareas, estar en otro lugar, haciéndoles reír, ayudándoles a hacerse preguntas, pero también acompañándoles y aconsejándoles en su travesía por la vida.

Abro el periódico y leo que la carrera más solicitada este año por los estudiantes es filosofía. Nuestra vida entera está volcada en las máquinas, los ordenadores tienen mayor computación que nuestro cerebro, y el *big data* se realiza de manera automática a través de la inteligencia artificial de esas máquinas. Todo es posible, todo es factible, lo importante no es el qué podemos hacer si no el por qué y el para qué, la filosofía, la ética, el hacer las cosas sostenibles, entender como nuestras acciones y la de las máquinas afectan a nuestro presente y futuro, el de los otros y lo que nos rodea —y volvemos a buscar nuevos sentidos a la vida. Las humanidades son protagonistas de nuevo, y a través de ellas, volvemos a acercarnos a aquello que nos hace únicos como humanos.

Esto es sólo un pequeño esbozo de lo que nos depara el futuro. Un futuro que se acelera exponencialmente debido a los grandes avances tecnológicos, el mundo de los datos, y la IA que cambiarán el mundo tal y como lo conocemos. No nos faltan desafíos. Hay mil millones de personas que quieren tener electricidad, millones carecen de agua potable, el clima está cambiando, la fabricación es ineficiente, el tráfico ahoga las ciudades, la educación es un lujo para muchos, y la demencia o las enfermedades degenerativas nos afectarán a casi todos si vivimos lo suficiente. No obstante, los científicos y tecnólogos, los emprendedores, y el mejor entendimiento de la sociedad, conjuntamente con la gran revolución de la IA nos permitirá encontrar soluciones para muchos de estos problemas; y estoy seguro de que lo conseguiremos. Prepárate para el nuevo mundo que está por llegar.

La IA en sí misma todavía está en una etapa incipiente. Gracias a los avances de los últimos años, estamos empezando a construir sistemas que pueden percibir, aprender y razonar, y sobre esta base, pueden hacer predicciones o recomendaciones. Casi todos los campos de la actividad humana podrían beneficiarse de los sistemas de IA diseñados para complementar la inteligencia humana. Desde la prevención de enfermedades que alguna vez fueron mortales, hasta permitir que las personas con discapacidades participen más plenamente en la sociedad, hasta la creación de formas más sostenibles de utilizar los escasos recursos de la tierra. La IA promete un futuro mejor para todos.

Un cambio de esta magnitud inevitablemente origina problemas sociales. La era de los datos y la informática nos ha obligado a lidiar con preguntas importantes sobre la privacidad, la seguridad, la equidad, la inclusión y la importancia y el valor del trabajo humano. Todas estas preguntas cobrarán una importancia particular a medida que los sistemas IA se vuelvan más útiles y se desplieguen más ampliamente.

Esto tomará un enfoque centrado en el ser humano. Y tomará un enfoque que esté centrado en el aprovechamiento del poder de la inteligencia artificial para ayudar a

las personas. La idea no es reemplazar a las personas con máquinas, sino complementar las capacidades humanas con la capacidad incomparable de la IA para analizar grandes cantidades de datos y patrones que de otro modo serían imposibles de detectar.

Cómo la IA cambiará nuestras vidas, y la vida de nuestros hijos, es imposible de predecir. Pero seguro que ayudará a muchas personas jóvenes y mayores que tienen ideas imaginativas sobre cómo utilizar la inteligencia artificial para abordar los desafíos sociales. Un enfoque centrado en el ser humano sólo puede realizarse si los investigadores, los responsables de la formulación de políticas y los líderes del gobierno, las empresas y la sociedad civil se unen para desarrollar un marco ético compartido para la inteligencia artificial.

Esto, a su vez, ayudará a fomentar el desarrollo responsable de los sistemas de inteligencia artificial que engendrarán confianza para sentar las bases de una IA centrada en el ser humano en la que todos confíen.

Bibliografía

CICI, BLERIM, MARKOPOULOU, ATHINA, FRÍAS-MARTÍNEZ, ENRIQUE, y LAOUTARIS, NIKOLAOS, «Assessing the potential of ride-sharing using mobile and social data: a tale of four cities», Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

LLORENTE, A., GARCÍA-HERRANZ, M., CEBRIÁN, M. y MORO, E. «Social media fingerprints of unemployment», Plos One, 2015.

PENTLAND, ALEX, *Social Physics: How Good Ideas Spread — The Lessons From A New Science*, The Penguin Press, Londres, 2014.

SAN PEDRO, JOSÉ, PROSERPIO, DAVIDE y OLIVER MOBISCORE, NURIA, «Towards Universal Credit Scoring from Mobile Phone Data». Conferencia en la UMAP 2015. Publicada en *Lecture Notes in Computer Science*, Vol. 9.146, pp. 195-207. Dublín, 2015.

TOOLE, J. L., LIN, Y. R., MUEHLEGGER, E., SHOAG, D., GONZÁLEZ, M.C., y LAZER, D., «Tracking employment shocks using mobile phone data», *Journal of The Royal Society Interface*, 12 (107), 2015.

<<https://www.partnershiponai.org>>.

<<https://www.nytimes.com/2017/10/26/opinion/algorithm-compass-sentencing-bias.html>>

<<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>>.

<<https://www.nytimes.com/2017/11/21/magazine/can-ai-be-taught-to-explain-itself.html>>.

<<https://www.linkedin.com/pulse/today-technology-day-horse-lost-its-job-brad-smith>>.

<http://www3.weforum.org/docs/WEF_FOJ_Executive_Summary_Jobs.pdf>.

<<http://query.nytimes.com/gst/abstract.html?res=9C03EEDF1F39E133A25755C2A9649C946995D6CF&legacy=true>>.

<<https://www.economist.com/news/special-report/21700758-will-smarter-machines-cause-mass-unemployment-automation-and-anxiety>>.

<<https://www.economist.com/news/special-report/21700758-will-smarter-machines-cause-mass-unemployment-automation-and-anxiety>>.

<<https://www.economist.com/news/special-report/21700758-will-smarter-machines-cause-mass-unemployment-automation-and-anxiety>>.

<<https://venturebeat.com/2017/10/04/the-fundamental-differences-between-automation-and-ai>>.

<https://www.washingtonpost.com/news/the-worldpost/wp/2017/10/19/inside-chinas-quest-to-become-the-global-leader-in-ai/?utm_term=.9da300d7d549>.

<https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment>.

<<https://openknowledge.worldbank.org/handle/10986/23347>>.

<https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2940245>.

<<https://www.theguardian.com/technology/2017/jan/11/robots-jobs-employees-artificial-intelligence>>.

<https://www.postandcourier.com/business/as-amazon-pushes-forward-with-robots-workers-nd-new-roles/article_c5777048-97ca-11e7-955e-8f628022e7cc.html>.

<<https://www.forrester.com/report/The+Future+Of+Jobs+2025+Working+Side+By+Side/E-RES119861>>.

<<https://www.economist.com/news/special-report/21700758-will-smarter-machines-cause-mass-unemployment-automation-and-anxiety>>.

<http://reports.weforum.org/future-of-jobs-2016/skills-stability/doing_wp_cron=1514488681.1306788921356201171875>.

<<https://www.technologyreview.com/s/515926/how-technology-is-destroying-jobs>>.

<<https://cew.georgetown.edu/wp-content/uploads/Americas-Divided-Recovery-web.pdf>>.

<https://krueger.princeton.edu/sites/default/files/akrueger/les/katz_krueger_cws_-_march_29_20165.pdf>.

<<http://www.oxfordmartin.ox.ac.uk/publications/view/1314.36>.
http://www.hamiltonproject.org/papers/who_is_out_of_the_labor_force>.

<<http://www.pewinternet.org/2016/11/17/gig-work-online-selling-and-home-sharing>>.

<<http://www.bloomberg.com/news/articles/2015-08-18/why-6-million-americans-would-rather-work-part-time>>.

<<https://www.teacherspayteachers.com>>.

<<http://journals.sagepub.com/eprint/3FMTvCNPJ4SkhW9tgpWP/full>>.

<<http://globalworkplaceanalytics.com/resources/costs-benets>>.

<<http://www.pewsocialtrends.org/2016/10/06/4-skills-and-training-needed-to-compete-in-todays-economy>>.

<<https://secure-media.collegeboard.org/digitalServices/pdf/research/2016/Program-Summary-Report-2016.pdf>>.

<<https://www.bls.gov/charts/job-openings-and-labor-turnover/opening-hire-seps-rates.htm>>.

<<https://www.bloomberg.com/news/articles/2017-06-22/the-world-s-workers-have-bigger-problems-than-a-robot-apocalypse>>.

<<https://www.nationalskillscoalition.org/resources/publications/2017-middle-skills-fact-sheets/le/United-States-MiddleSkills.pdf>>.

<http://burning-glass.com/wp-content/uploads/2015/06/Digital_Skills_Gap.pdf>.

<<https://www.nationalskillscoalition.org/resources/publications/le/Opportunity-Knocks-How-expanding-the-Work-Opportunity-Tax-Credit-could-grow-businesses-help-low-skill-workers-and-close-the-skills-gap.pdf>>.

<<https://news.microsoft.com/2017/06/27/the-markle-founda-tion-and-microsoft-partner-to-accelerate-a-skills-based-labor-mar-ket-for-the-digital-economy>>.

<https://www.weforum.org/agenda/2015/02/a-brief-history-of-big-data-everyone-should-read/>

<http://laoutaris.info/index.php/category/dtl/>

http://chutzpah.typepad.com/slow_movement/page/21/

<https://www.cnn.com/2017/03/13/tim-berners-lee-we-must-tackle-fake-news-for-the-benefit-of-all-of-humanity.html>

https://en.wikipedia.org/wiki/Flash_Boys

<https://www.nytimes.com/2010/05/02/magazine/02self-measurement-t.html>

<https://www.cnn.com/2015/08/20/travel/delta-lightning-strike-atlanta-airport-feat/index.html>

<https://thefloatinglibrary.com/2010/05/09/self-tracking-is-the-future/>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4528087/>

<https://www.frontiersin.org/articles/10.3389/fpubh.2015.00189/full>

<https://www.iotacademy.ir/en/articles/internet-things-far-bigger-anyone-realizes>

https://www.huffingtonpost.com/thought-matters/algorithms-to-live-by_b_9772622.html

<https://medium.com/owkin/transfer-learning-and-the-rise-of-collaborative-artificial-intelligence-41f9e2950657>

<https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>

<http://www.gizmodo.co.uk/2015/02/the-ai-revolution-how-far-away-are-our-robot-overlords/>

<https://www.inverse.com/article/7852-ai-revolution-road-superintelligence>

<https://heleo.com/tim-urban-what-is-the-future-of-ai/5544/>

<https://www.theguardian.com/odine-partner-zone/2017/apr/03/who-really-owns-data>

<https://dzone.com/articles/how-data-is-changing-the-food-you-eat>

https://www.huffingtonpost.in/entry/from-big-data-to-artifici_b_10817892

<http://weku.fm/post/remembering-when-driverless-elevators-drew-skepticism>

<http://earth4all.net/ai-the-most-important-topic-of-human-future/>

<https://www.weforum.org/agenda/2016/10/how-google-is-changing-our-brains>

http://www.stltoday.com/business/local/why-a-staggering-number-of-americans-have-stopped-using-the/article_e83e8ccf-115b-5836-8a8f-c5a821e66d76.html

<https://www.livetradingnews.com/privacy-security-concern-chilling-internet-users-4622.html>

<https://www.iotcentral.io/blog/how-big-data-is-changing-the-world-of-soccer>

<https://www.nytimes.com/2012/12/09/business/company-envisions-vaults-for-personal-data.html>

<https://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource>

<https://blog.frontiersin.org/2016/06/07/you-are-not-as-anonymous-as-you-think-online/>

https://alexgkendall.com/artificial_intelligence/lets_talk_about_ethics_in_artificial_in

<http://www.conversationagent.com/2016/09/thirty-seven-percent.html>

<http://www.themoviespoiler.com/2014Spoilers/ImitationGame.html>

<https://uk.news.yahoo.com/food-scientists-discovered-surprising-principle-154700667.html>

<https://ai.xprize.org/news/people-centered-data-privacy-ai-world>

<https://www.fnac.pt/The-Master-Algorithm-Pedro-Domingos/a897677>

<https://squattheplanet.com/threads/most-connected-man-in-the-world.25031/>

https://infogalactic.com/info/Pareto_principle

https://en.wikipedia.org/wiki/Stanley_Milgram

Agradecimientos

Hoy vivimos en un futuro que se acelera debido a los grandes avances tecnológicos que cambiarán el mundo tal y como lo conocemos. No nos faltan desafíos. Hay mil millones de personas que quieren tener electricidad, millones de personas carecen de agua potable, el clima está cambiando, la fabricación es ineficiente, el tráfico ahoga las ciudades, la educación es un lujo, y la demencia o otras enfermedades degenerativas nos afectarán a casi todos si vivimos lo suficiente.

Y sin embargo, científicos, ingenieros y tecnólogos trabajan infatigablemente, se niegan a rendirse y tratan de encontrar soluciones; y estoy seguro de que lo conseguirán. Mi agradecimiento a todos ellos.

En un mundo cambiante, donde la única manera de predecir el futuro es inventándolo, empecé mi carrera como ingeniero de telecomunicaciones en la Universidad Pública de Navarra (UPNA). Esta universidad consiguió imprimir en mí y en generaciones venideras de ingenieros de telecomunicaciones las herramientas necesarias para poder desarrollar un espíritu de investigación y emprendeduría.

En el Politécnico Federal de Lausanne, donde me formé como científico, aprendí a hacer algo nuevo cada día que me diese un poco de miedo, a asumir riesgos y a no mirar atrás. Jean Yves Le Boudec, Jean Pierre Hubaux, Martin Vetterli y, cómo no, mi director de tesis, Ernst Biersack, han sido grandes profesores que me enseñaron lo que hoy sé sobre el mundo de Internet, las redes y la ciencia de la computación.

Desde mi tarea como profesor en la universidad de Columbia en Nueva York, ingeniero en el Silicon Valley, investigador en Bell-Labs o Microsoft Cambridge, a mi actual responsabilidad en Telefónica, siempre he tenido presente los consejos que recibí de todos mis mentores y el trampolín que me ofrecieron para pensar libremente, de forma crítica y sin barreras. Todo mi agradecimiento a Alfonso Carlosena, Ken Morse, Paul Francis, Christophe Diot, Balachander Krishnamurthy, Sandeep Sibal, Sanjoy Paul, Krishan Sabnani, Vitali Fridman, Henry Chesbrough, y tantos otros que me han guiado, y aconsejado, y que crean cada día un valor diferencial de innovación para toda la sociedad.

Desde mi primer contacto con un sistema UNIX como becario del departamento de física en el King's College, hasta mi último trabajo de investigación, he tenido la suerte de trabajar con grandes colegas profesionales de las que he aprendido enormemente. Ant Rowstron, Christos Gkantsidis, Dina Papagiannaki, Nikos

Laoutaris, Thomas Karagiannis, Vishal Misra, Nuria Oliver, están explorando las redes, la inteligencia artificial, y conectando el planeta con su visión, conocimiento y fuerza creativa.

La vista siempre puesta en el futuro me ha empujado a una carrera profesional en el mundo de la innovación que en ningún momento me podía imaginar. Carlos Domingo, David del Val, José María Álvarez Pallete, Ferran Adrià, Ian Small, Carlos Colomer, Gonzalo Martín-Villa, Oriol Lloret, Xavier Capellades, Esther Gravalos, Kim Faura, Bruno Vilarasau, Ramon Salabert, Ramón Sanguesa, Raúl Ortega, Vicente San Miguel, Julio Linares, el equipo de Telefónica I+D, el equipo de Synergic Partners, Luca, Aura, y la cuarta plataforma, y tantos otros que me han inspirado en este viaje de descubrimiento y impacto. Gracias.

Gracias, Montse, por tus comentarios, correcciones y siempre buena y contagiosa energía. Roger Domingo, por empujarme a escribir este libro. Andrea, por tu paciencia y cuidados. Brian, por tu creatividad, consejos y una buena historia de amistad. Y a todo el equipo de Alpha, por traer más esperanza a este mundo.

Natalia, gracias por ayudarme siempre en lo que he creído y a que lo imaginario se convierta en real, y por el regalo de nuestro hijo Uriel.

Antonio, Raquel, Adriana y Carlota, por ser mi familia.

Y por último, gracias también, y como no podría ser de otra forma, a mis padres, sin cuya guía, apoyo y comprensión no hubiera llegado hasta aquí. A ellos va el mayor de los reconocimientos.



PABLO RODRÍGUEZ RODRÍGUEZ (Oviedo, Asturias, España, 17 de abril de 1972). Científico de la computación e investigador español. Es muy conocido por sus estudios sobre el compartir archivos Peer-to-Peer y por el contenido generado por el usuario (user-generated content).

Hizo su pregrado y máster en ingeniería de telecomunicaciones en la Universidad Pública de Navarra, continuo sus estudios en el King College de Londres en física computacional y allí realizó estudio sobre sensores electro-ópticos y, luego en Francia y Suiza obtuvo un doctorado en ciencias de la computación. Ha escrito muchos artículos en ciencias de la computación y su el libro *Inteligencia artificial: como cambiará el mundo (y tu vida)*, 2018.