

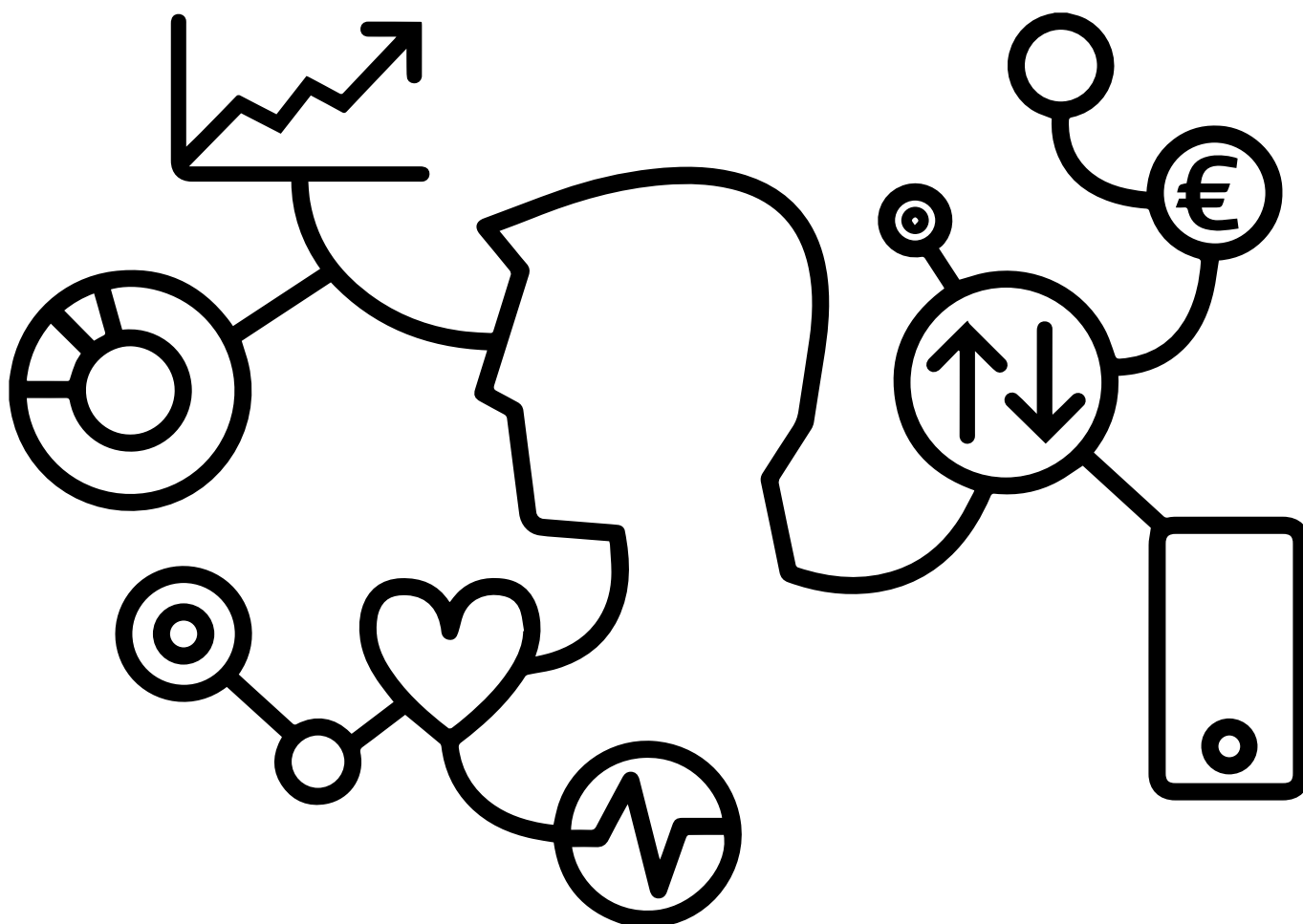


- > **Nuestra producción de datos ha crecido exponencialmente en pocos años.**
Por qué son necesarias herramientas que analicen y procesen la información relevante.
- > **Las implicaciones del *Big Data* para cada uno de nosotros como individuos y nuestras relaciones con el entorno.**
- > **Cómo mejorar nuestro rendimiento humano con el procesamiento de los grandes volúmenes de datos.**

Big data

El poder de los datos

"La tecnología big data ya está aquí, ahora hay que aprender a usarla."



Big data

El poder de los datos

Agradecimientos

Nuestro agradecimiento a **Esther Paniagua**, periodista y autora de este informe. Su trabajo ha sido decisivo para poder plasmar las conclusiones de la esta tendencia del Future Trends Forum.

Nuestro agradecimiento a todos los miembros del Future Trends Forum (FTF) que han hecho posible el éxito de nuestra última reunión, especialmente a aquéllos que han participado activamente en la realización de esta producción:

Por su inestimable colaboración en la elaboración de esta publicación:

Chris Meyer
Lita Sands
Vicki Seyfert-Margolis
Stefan Bungart
J.C. Herz

En la organización y metodología de la reunión del Future Trends Forum:

Chris Meyer

Y por último, agradecer a las personas del equipo, por su compromiso y buen hacer en el desarrollo del contenido de esta publicación:

Fundación Innovación Bankinter
Sergio Martínez-Cava
Marce Cancho
María Teresa Jiménez
Lara García de Vinuesa
Dorsey Lockhart
Pablo Lancry



Garrick Jones
Clemens Hackl



Raquel Durán

Las opiniones expresadas en este informe son del autor y no reflejan la opinión de los expertos que participaron en la reunión del Future Trends Forum.

Ponentes y asistentes

Ana María Arboleda

Directora General de Atlantis Healthcare Spain

Miguel Arias

COO en CartoDB

Puneet Batra

Cofundador de LevelTrigger

Jessica Bland

Investigadora de Tecnologías del Futuro de Nesta

Carla Brodley

Decana de la College of Computer and Information Science de la Northeastern University

Stefan Bungart

Líder de GE Software Europe en GE Global Research

Antonio Damasio

Profesor y Director del Brain and Creativity Institute en la Universidad del Sur de California (USC) y Premio Príncipe de Asturias de Investigación Científica y Técnica

Hanna Damasio

Directora del Dana and David Dornsife Cognitive Neuroscience Imaging Center de la Universidad del Sur de California (USC) y Premio Príncipe de Asturias de Investigación Científica y Técnica

Maarten den Braber

Co-fundador de SHIFT

Julie Freeman

Artista y líder de Datos del Culture Art Programme en el Open Data Institute (ODI)

J.C. Herz

Diseñadora y consultora estratégica con Big Data

Eduardo Hruschka

Jefe Científico de Datos en Big Data Brazil

Richard Kivel

Emprendedor e inversor en Tecnología médica y biofarmacia y Patrono de Fundación Innovación Bankinter

Joseph Kvedar

Director del Center for Connected Health

Gregory La Blanc

Investigador Haas Economic Analysis en la Universidad de Berkeley

Philip Lader

Presidente no ejecutivo de WPP Group y Patrono de Fundación Innovación Bankinter

Julia Li

Fundadora y CEO de HCD Global

Juan Carlos López

Líder de Relaciones Académicas y colaboración de Roche Innovation Center

Chris Meyer

CEO de Nerve LLC y Patrono de Fundación Innovación Bankinter

Emilio Méndez

Director del Center for Functional Nanomaterials del U.S. Department of Energy's, Brookhaven National Laboratory, Premio Príncipe de Asturias 1998 y Patrono de Fundación Innovación Bankinter

Esteban Moro

Investigador en la Universidad Carlos III de Madrid

Iyad Rahwan

Profesor de Tecnologías de la Información en el Masdar Institute

Pablo Rodríguez

Investigador y Director de Innovación en Telefónica

Lita Sands

Líder global de Transformación Digital de Novartis

Vicki Seyfert-Margolis

Fundadora y CEO de My Own Med, Inc.

Eden Shochat

Fundador de Aleph y Patrono de Fundación Innovación Bankinter

Taryn Sullivan

CEO de Efficiency Exchange

Steve Trachtenberg

Presidente Emérito de la George Washington University y Patrono de Fundación Innovación Bankinter

Wilfried Vanhonacker

Profesor en Olayan School of Business, AUB y Patrono de Fundación Innovación Bankinter

Andreu Veà

Presidente de la Internet Society (ISOC-ES)

Muchas gracias,
Fundación Innovación Bankinter

Índice

008	Chris Meyer Prólogo		8/9
010	Cap. 1 Introducción		11/12
015	Cap. 2 Show me the money	2.1 Cuantificar 2.2 Predecir y cambiar 2.3 Optimizar 2.4 Innovar para transformar	19/20 21/23 24/26 27/30
032	Cap. 3 De personas para personas	3.1 Gobernantes y ciudadanos 3.2 Médicos y pacientes 3.3 Directivos y empleados 3.4 Entrenadores y jugadores 3.5 Policías versus criminales	37/38 39/40 41/42 43/44 45/46
048	Cap. 4 Los retos pendientes	4.1 Qué medir y cómo librarse del ruido 4.2 Entender y aplicar 4.3 Propiedad versus privacidad versus seguridad	53/55 56/58 59/61
062	Cap. 5 La llave al futuro	5.1 Legislación y marco legal 5.2 Silos 5.3 Formatos, estándares y calidad 5.4 Factor humano y resistencia 5.5 Habilidades y recursos 5.6 Mercados infradesarrollados	66/67 68 69 70 71 72
073	Infografía El futuro del Big Data	Time line hitos destacados	74/75
076	Cap. 6 Glosario		76/78
079	Cap. 7 Recursos	7.1 Informes y estudios 7.2 Artículos y noticias 7.3 Sitios web 7.4 Aplicaciones y plataformas 7.5 Otros	79 79 80 80 80



Big data in context

Chris Meyer. Fundador de Monitor Talent y Patrono de Fundación Innovación Bankinter

➔ **En la década de los 90,** cuando la informatización de la cada una de las funciones de negocio empezó a dar paso a los paquetes de software para empresas, los profesionales de la informática jerarquizaban la utilidad en "datos, información, conocimiento, sabiduría".

Los datos eran materia prima que había que agregar para generar información; los patrones que entonces se perfilaban llevaban al conocimiento, que a su vez nos daría pie para llegar a ideas esclarecedoras y nos haría más sabios. El software corporativo que recababa datos de todas las áreas de una organización en "almacenes de datos" (*data warehouses*) se analizaba periódicamente para descubrir las ideas más esclarecedoras respecto al rendimiento del negocio.

Pero por el camino hubo un giro inesperado: internet. La red de redes adelantó por la izquierda a los datos corporativos y encontró un atajo hacia el conocimiento. En los últimos veinte años, gran parte del conocimiento acumulado en el mundo se ha hecho accesible. En cuanto se genera conocimiento nuevo (artículos médicos, guías de viaje, críticas de la cultura pop), se hace público. A medida que el valor de acceder a internet se ha disparado, el aliciente para granjearse dicho acceso ha subido a la par, y

los teléfonos móviles inteligentes se están haciendo universales, haciendo llegar el conocimiento a prácticamente todos los lugares (el túnel del metro y el campo incluidos).

Por tanto, el desarrollo de la economía de la información está dando pasos atrás para tomar un camino que había dejado de lado. Cada dispositivo conectado es también un sensor capaz de recabar datos sobre la ubicación de las personas, el rendimiento de las máquinas, los contaminantes de las fábricas, la humedad de los campos, la salud de los pacientes y los sentimientos de las personas –según publican ellas en las redes sociales o se observa en su comportamiento. Esta proliferación de sensores ha generado el torrente de *Big Data* –cantidades ingentes de tomas de datos, muchas en tiempo real, que de momento son más rápidas que la capacidad de agregarlas al nivel de conocimiento. Pero la industria va a toda velocidad para ponerse a la altura.

En esta publicación leerá sobre la explosión de *Big Data*, un término que abarca:

- > **Una creciente concienciación** entre ejecutivos, políticos e individuos sobre la nueva disponibilidad de los datos para ayudar en la toma de decisiones. Pongamos por caso un sistema de navegación en automóvil que propone una ruta alternativa basada en información del tráfico en tiempo real, generada a partir de los datos de viaje de miles de teléfonos móviles con GPS incorporado. O los sensores en los contadores de la electricidad, que ayudan a las eléctricas a gestionar los picos de carga.
- > **Una industria en plena explosión** que se ofrece a dar buena cuenta de *Big Data* a través de una infraestructura técnica que recoge, almacena y da acceso a los datos en sí. El acceso continuo ha sustituido el modelo de data *warehouse* centrado en los lotes (*batch*) – los sistemas urbanos de transporte actualizan en tiempo real cuándo llegará el autobús o el tren. Los datos acumulados se transforman rápidamente en información sobre el rendimiento del sistema.

- > **Y en la vanguardia**, las técnicas aplicadas de la ciencia de datos crean conocimiento y dan incluso ideas esclarecedoras a partir de esos datos. Por ejemplo, un investigador médico recabó una amplia base de datos de historiales médicos y los cargó en un algoritmo de aprendizaje computacional sin ningún objetivo en concreto; el programa descubrió que la diabetes tipo II no es una, sino cuatro enfermedades distintas.

Y una conjetura: el neurólogo Antonio Damasio (que participó en Future Trends Forum) ha escrito que la percepción se desarrolló en los humanos con la conciencia de nuestro estado interior para integrar datos sobre el mundo exterior recabados a partir de los cinco sentidos. *Big Data* está añadiendo la conciencia exterior a nuestros sistemas de información, que en principio estaban contruidos para describir el estado interno de las organizaciones.

Cuando los *smart phones*, de forma espontánea, nos digan que hay que irse porque está empeorando el tráfico que nos separa del evento para el que compramos las entradas por internet la semana pasada, ¿diremos que los móviles son conscientes?

Big Data está de moda ahora mismo porque ofrece recompensas generosas e inmediatas y motiva las decisiones en tiempo real, como el precio de una habitación de hotel para un cliente dado, la decisión de reponer inventario o buscar ciertos recursos de marketing. Pero igual que con cada avance en la economía de la información, el desarrollo de recursos técnicos y humanos para extraer valor de *Big Data* llevará a su vez al siguiente paso, en el cual los sistemas de información generen habitualmente conocimiento e ideas esclarecedoras. El informe a continuación describe los últimos avances, los retos de cara al futuro, y da unas pinceladas a lo que podría ser posible cuando nos pongamos en marcha.



Introducción

1

Nuestra capacidad para producir información ha avanzado a la velocidad de la luz con respecto a hace tan solo unos años. La ingente cantidad de datos disponibles ha hecho necesario el desarrollo de herramientas que los analicen y procesen para identificar lo relevante. Estas tecnologías ya existen pero apenas hemos empezado a exprimir sus posibilidades.

Esther Paniagua

Periodista y autora del informe FTF

En los últimos cinco años se ha generado más información científica que en toda la historia de la humanidad.

"Si se pusieran en fila todos los nuevos libros publicados, nos deberíamos desplazar a 150 kilómetros por hora para mantenernos al frente de la hilera", escribía Stephen Hawking en *El universo en una cáscara de nuez* hace ya trece años. "Si continuara el crecimiento exponencial, se publicarían diez artículos por segundo en mi especialidad de física teórica, y no tendría tiempo de leerlos", proseguía el aclamado físico.

En efecto, eso ha sido exactamente lo que ha pasado. Según el profesor de bioinformática de la Escuela de Salud Pública de Harvard Winston Hide, **en los últimos cinco años se ha generado más información científica que en toda la historia de la humanidad.** El volumen de los datos existentes es de tal magnitud que, si ocupara un espacio físico, Hide asegura que superaría el tamaño de una galaxia.

Esta aceleración de la velocidad a la que generamos contenido ha propiciado la emergencia de una nueva ciencia: el *big data*, que en español se podría traducir como "datos masivos". Su significado parece claro, sin embargo, este concepto no es suficiente per se para explicar el fenómeno, que trasciende a una simple idea de cantidad. No estamos hablando solo de un montón de datos, sino de la generación de valor a partir de su procesamiento y análisis. Esa es la verdadera **revolución del big data** y el motivo que ha llevado a una treintena de expertos de todo el mundo a reunirse en la XXIII edición del *Future Trends Forum* organizado por la Fundación Innovación Bankinter, que ha inspirado esta publicación.

Para llegar hasta aquí se han desarrollado numerosos dispositivos y aplicaciones que permiten medir, estructurar, procesar y analizar grandes volúmenes de datos gracias a nuevas infraestructuras tecnológicas diseñadas para soportar *petabytes* de información. Constantemente aparecen nuevas y refinadas tecnologías que están permitiendo mejorar la eficiencia de las organizaciones, precisar los gustos y necesidades de clientes y consumidores, conectar a Gobiernos y ciudadanos, reforzar la seguridad y la inteligencia de las ciudades, realizar predicciones de todo tipo y generar **patrones de comportamiento** útiles para que las empresas conozcan mejor a su público y para concienciar a las personas sobre sus propios hábitos, incluso dando recomendaciones de cómo cambiarlos para mejorar aspectos de su vida como la salud.

Todo esto es solo el comienzo, ya que apenas estamos empezando a aprovechar el potencial que ofrece el análisis de datos masivos. El *big data* ha

abierto la veda a una sociedad más informada, más eficiente, capaz de realizar proezas hasta hace poco reservadas al ámbito de la ciencia ficción.

Tal vez Hawking no tenga tiempo de leer todos los artículos científicos que se generen, pero tampoco le hará falta. Alguna tecnología *big data* lo hará por él, y solo tendrá que preocuparse de interpretar los resultados. De hecho, no tendrá que esperar mucho: trece años después de sus vaticinios, esto ya forma parte del presente. **La toma consciente de decisiones** basadas en datos es hoy una realidad, y quien no la aplique se quedará atrás.

El reto ahora es encontrar la aguja en el pajar, identificar **lo relevante** y aplicarlo involucrando a todos los agentes implicados (empresas, organizaciones, gobiernos, profesionales y ciudadanos), superando las barreras legales, estructurales y humanas que presenta. Nadie dijo que fuera fácil, pero merece la pena intentarlo.

El big data ha
abierto la veda a
una sociedad más
informada, más
eficiente, capaz de
realizar proezas
hasta hace poco
reservadas al ámbito
de la ciencia ficción.



Show me the money

2

Cap. 2/7

2.1 Cuantificar

2.2 Predecir y cambiar

2.3 Optimizar

2.4 Innovar para transformar



Convertir los datos en oro

Lita Sands. Líder global de Transformación Digital de Novartis

➔ **Sin duda,** hemos abierto la caja de Pandora, y dado que las tecnologías digitales siguen transformando prácticamente todos los aspectos de la vida, es excepcional que una gran empresa pueda permitirse quedarse en el banquillo. Estamos viviendo una revolución en cómo nos relacionamos con el mundo a nuestro alrededor, y los datos consiguientes están transformando con celeridad modelos de negocio aparatosos.

"El 90% de todos los datos que generan dispositivos como *smartphones*, tabletas, vehículos y electrodomésticos conectados nunca se analiza... el 60% de estos datos empiezan a perder valor en cuestión de milisegundos..." **IBM**¹

"Todos los días, creamos 2,5 trillones de bytes en datos... el 90% de los datos del mundo a día de hoy se han creado en los últimos dos años." **SINTEF**²

"Datos, datos por todas partes pero ni un solo pensamiento en que reflexionar."
Theodore Roszak

Pongamos por caso el mercado de los seguros de vida. John Hancock Financiamiento anunció hace poco que iba a ofrecer descuentos a asegurados que lleven puestos sistemas para monitorear la forma física conectados a Internet. Con cada nuevo seguro de vida, se ofrece una pulsera Fitbit para monitorear la actividad física. Cuanto más ejercicio se haga, más se ahorra (hasta un 15%)³. Literalmente de la noche a la mañana, han creado una ventaja competitiva y además están gestionando un grupo de clientes más sano.

Pongamos por caso el descubrimiento de fármacos en la industria farmacéutica. Watson de IBM, en asociación con la facultad de medicina de Baylor, ha descubierto seis proteínas que modifican la p53—una proteína fundamental en muchos casos de cáncer—, lo cual les ha llevado a descubrir un fármaco nuevo en cuestión de semanas. Esto es destacado, teniendo en cuenta que en los últimos 30 años, la media en la comunidad científica ha sido de un descubrimiento de una proteína diana de este tipo al año⁴.

Todos y cada uno de los aspectos de nuestra vida dejan una huella digital que se puede analizar y podría redundar en una idea esclarecedora. Microsoft supo trastear entre datos de búsqueda para identificar los primeros indicios de efectos secundarios nocivos en medicamentos que nunca antes se habían detectado⁵.

La transparencia que esto introduce puede ser esclarecedora y abrumadora tanto para clientes como empleados. Al sustituir los folletos en papel por iPads, las empresas farmacéuticas aprovechan los datos de marketing en un círculo cerrado para crear indicadores tempranos de ventas con una precisión milimétrica, pudiendo llegar hasta el emisor de la receta si lo quisieran. Compare esto con lo que existía antes: o no tener ni la más mínima idea porque el acceso a esos datos era cero, o la compra de datos muy caros y retrasados que no se podían atribuir al rendimiento de ninguna persona en particular.

Imagínense la reacción del comercial de ventas al que ahora se puede responsabilizar del impacto directo de su presentación. O el director de marketing de marca,

al que se le medía por actividad pero ahora se le mide por resultados.

Empleados en todas las grandes industrias están viviendo esta ola de cambio. Los datos traen ideas esclarecedoras, transparencia y responsabilización. Lo digital trae la posibilidad de una mano de obra colaboradora y conectada que trabaja más rápido, mejor, de forma más inteligente, aunque esté repartida por todo el mundo. Pero pocas empresas están preparadas para lo que va a ser el renacimiento y la redefinición de la empresa del siglo XXI.

Este capítulo entrará en estos temas, y en cómo debemos tener en cuenta el qué y el cómo en la transformación para competir con eficacia en este nuevo mundo.

1: ibm.biz/BdXRSt

2: **SINTEF**. "Big Data, for better or worse: 90% of world's data generated over last two years." ScienceDaily. ScienceDaily, 22 May 2013. www.sciencedaily.com/releases/2013/05/130522085217.htm.

3: http://www.johnhancock.com/about/news_details.php?fn=apr0815-text&yr=2015#.VTPAjpftvsU_twitter

4: ibm.biz/BdFYLE

5: <http://research.microsoft.com/en-us/um/people/horvitz/pharmacovigilance-signals%20from%20the%20crowd.pdf>

Introducción

2.0

➔ **Ventaja competitiva.** Son las dos palabras mágicas que ha pronunciado el big data.

Su promesa: generar nuevas oportunidades de negocio a partir de datos de alto valor añadido. Su herramienta: las nuevas plataformas y soluciones para registrar y procesar interacciones y actividades de negocio. Su poder: la capacidad para revelar nueva información sobre las cadenas de valor -de la organización y del sector- y para abordar problemas de negocio antes irresolubles. Compañías como Google o Amazon lo han visto muy claro desde el principio. No en vano, su modelo de negocio pasa por la explotación de los valiosos datos que millones de usuarios depositan en sus plataformas. Pero no son las únicas. Hay compañías como Acxiom que viven de recopilar datos que generamos en nuestra vida personal *online* y de procesarlos para terceros que pagan por ellos. A juzgar por sus [ingresos de más de un billón de dólares anuales](#), no les va nada mal. Tampoco pueden quejarse empresas de software como Splunk, que con sus análisis de datos e informes visuales obtuvieron [más de 300 millones de dólares](#) en 2014.

"También hay compañías de telecomunicaciones vendiendo datos, otras que comercian con datos de Twitter y gente vendiendo información", señala el investigador y experto en *big data* del departamento



Esteban Moro

Investigador en la
Universidad Carlos III de Madrid

de Matemáticas de la Universidad Carlos III de Madrid **Esteban Moro**. Las opciones son múltiples y la realidad innegable es que hay un mercado para los datos, cuyo valor se empieza a apreciar. Hablamos del potencial de *big data* para transformar una industria, en lugar de simplemente hacerla más eficiente; de las oportunidades que ofrece a todos los niveles de toma de decisiones: organizacional, de gestión de procesos, de atracción al consumidor y de redefinir los modelos actuales de investigación, desarrollo de producto y comercialización; hablamos de un cambio transversal en áreas estratégicas y en cómo aplicarlo para cuantificar, predecir y cambiar, optimizar e innovar.

Cuantificar

2.1

En nuestra ropa, en nuestro móvil, en el reloj, en los semáforos y hasta en los contenedores. Están en todas partes y ni siquiera detectamos su presencia. Los sensores invaden nuestra realidad cotidiana, camuflados en todo tipo de objetos, aparatos y accesorios. Su función es recopilar información y son un elemento clave para el *big data*. Son, en muchos casos, el primer paso del proceso de la cuantificación. Y un paso nada desdeñable, por cierto.

Así lo cree **Joseph Kvedar**, fundador y director del Centro para la Salud Conectada de la Escuela Médica de Harvard, que destaca "el poder de la retroalimentación de los datos". Lo dice en referencia a aplicaciones de medición de actividad diaria -como [Moves](#)- que podemos descargar en nuestros *smartphones*, o a esas pulseras y relojes inteligentes -como [Fitbit](#) u [UP de Jawbone](#)- cada vez más presentes en nuestras muñecas y "que proporcionan datos objetivos y poderosos".

Estos datos, además, se pueden recopilar para obtener otros datos. Por ejemplo, sobre [cómo un terremoto ha afectado al sueño de las personas](#) en el área donde se ha producido. "Gracias a estas aplicaciones puedo darme cuenta de que gasto casi



Joseph Kvedar ▲
Director del Center for Connected Health

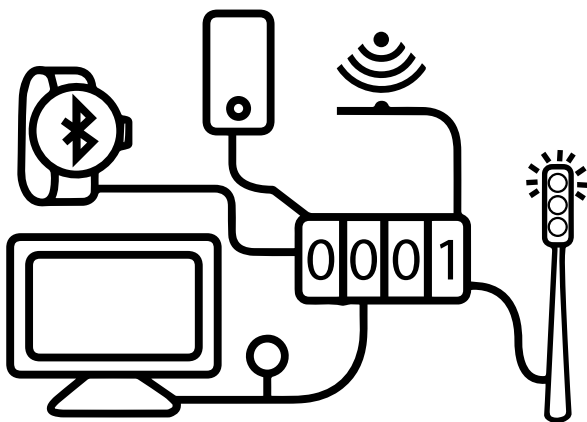


Maarten Den Braber ▲
Co-fundador de SHIFT

el doble de calorías retirando de mi jardín las hojas caídas del otoño que montando en bicicleta", afirma Kvedar. Es lo que ha bautizado como "**el poder del internet de las (saludables) cosas**". "Aumentan la concienciación y nos abren una oportunidad para aprender sobre nosotros mismos", añade.

Que se lo digan a **Maarten Den Braber**, cofundador de [SHIFT](#) y colaborador de la Singularity University. Den Braber es uno de los fundadores de la red en Europa de la comunidad [Quantified Self](#), creada para conocer las últimas novedades en herramientas y aplicaciones de automedición y para ayudar a la gente a entenderlas. Él mismo es un confeso adicto a **cuantificarse a sí mismo**. Muchos además comparten sus mediciones, tan curiosas como la [frecuencia cardíaca de un hombre mientras hace una propuesta matrimonio](#).

Aplicando este concepto de cuantificación en el entorno empresarial, y sin necesidad de sensores ni de modernos dispositivos de medición, hay todo un mundo de datos por cuantificar. No solo Google o Amazon disponen de información útil acerca de sus usuarios. El problema es que en la mayoría las organizaciones que los recogen no lo hacen con el fin de analizarlos sino simplemente por una



Los sensores invaden nuestra realidad cotidiana, camuflados en todo tipo de objetos, aparatos y accesorios.

cuestión de protocolos, que obligan a contar con registros pormenorizados y datos de todo tipo de los *stakeholders* con los que interactúan. Si bien estos datos están siendo recogidos y almacenados, a menudo no pasan de ahí, no llegan a ser procesados, ni analizados, ni utilizados con fines más allá de los meramente procedimentales ni, por supuesto, rentabilizados.

Esto es precisamente lo que está empezando a cambiar el *big data*: la necesidad de sacar partido de la masiva cantidad de información de la que las organizaciones son propietarias y la confluencia de esta necesidad con la aparición de herramientas que permiten hacerlo. Sectores como la banca o el negocio de las tarjetas de crédito y débito ya las utilizan. Estas empresas poseen datos valiosísimos acerca del comportamiento de los usuarios. "Las tarjetas de pago proporcionan un amplio abanico de datos acerca de consumidores y comerciantes", destacan expertos en el tema presentes en el FTF.

Según estos expertos, cada hora se pueden procesar una media de 60 millones de transacciones procedentes de cerca de 2.000 millones de tarjetas en 220 países y territorios, a través de más de 40 millones de comercios. "Con un conjunto de datos tan rico y amplio, podemos ser capaces de detectar cambios en la economía con una sensibilidad y profundidad increíbles", explican. Por eso, estos datos también sirven para ofrecer ideas a comerciantes, bancos, gobiernos y otros actores que les ayuden a tomar mejores decisiones acerca de cómo servir a sus públicos de interés.

Estos mismos expertos resumen los aspectos esenciales del *big data* en **cuatro V**:

- > **1. Volumen.** Es decir, el tamaño de los conjuntos de datos que generamos hoy en día. Para muestra un botón: los datos que producimos en el mundo a lo largo de dos días equivalen a todos los generados antes de 2003.
- > **2. Velocidad** (y frecuencia): Los datos se quedan desfasados y pierden su valor rápidamente. Sobre todo teniendo en cuenta que se generan datos cada segundo, incluidos entre ellos transacciones, fotos y video.
- > **3. Variedad:** Pueden proceder de sensores de calor, de cámaras, de nuestros *smartphones*, de nuestros pagos, de los coches, de sistemas de navegación y GPS, de las redes sociales, datos de viajes, postales y de otras fuentes.
- > **4. Veracidad:** La mayoría de los datos llegan en sucio, con campos que faltan o que son incorrectos. Esto puede ser más complejo si varios proveedores utilizan diferentes formatos y si los datos provienen de diferentes países, ya que pueden ser dramáticamente diferentes en función de las costumbres y usos locales. "Limpiar estos datos puede ser la actividad creadora más desafiante para generar valor", aseguran. El trabajo nunca acaba.

Según estos expertos,
cada hora se pueden procesar
una media de **60 millones** de
transacciones procedentes de cerca
de **2.000 millones** de tarjetas en
220 países y territorios,
a través de más de
40 millones de comercios

Predecir y cambiar

2.2

Predecir y cambiar comportamientos, que convierte al big data en una auténtica vía para la persuasión.

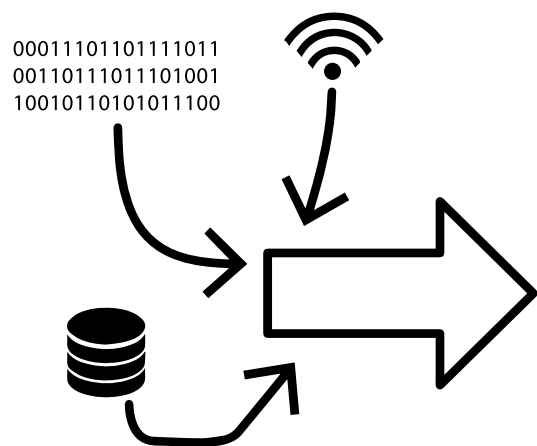
Pero, ¿para qué podemos usar todos estos datos? ¿De qué nos sirve cuantificar? Es imprescindible saber para qué lo estamos haciendo. Las posibilidades son múltiples y pueden ser muy ambiciosas. Una muy poderosa es la de predecir y cambiar comportamientos, que convierte al *big data* en una auténtica vía para la persuasión. La técnica requiere de refinamiento. Es necesario realizar un buen procesado y análisis de los datos cuantificados.

Como ejemplifican los participantes en el FTF, un buen trabajo permite apreciar datos tan "útiles e interesantes" como que en determinado periodo una gran cantidad de consumidores cambiaron sus patrones diarios de compra. Y no solo eso sino quiénes fueron, cuáles son sus características, en qué lugar y en qué condiciones lo hicieron. La utilidad del *big data* reside, según ellos, en descubrir "esas dos o tres cosas de las que no te habías dado cuenta y que salen a la luz gracias a los datos". "Es una gran oportunidad para obtener nuevos conocimientos sobre el comportamiento de los consumidores", sostienen. Y para predecirlos, porque **dichos conocimientos también nos permiten hacer predicciones.**

Los datos pueden proporcionar suficiente información como para pronosticar aspectos sobre los clientes, sobre los *stakeholders*, sobre la propia organización o el mercado en que opera. "Hace 15 años una cadena de restaurantes de marisco de Estados Unidos (EE. UU.) descubrió que los satélites del golfo de México transmitían datos sobre la temperatura del mar",

relata **Chris Meyer**. "Esto a su vez era un indicador de la fertilidad de las gambas, que crían a un nivel proporcional a la temperatura del agua en la que viven", prosigue. Así que los datos que proporcionaba el satélite servían a los propietarios de la cadena de restaurantes para predecir los precios a los que podrían adquirir las gambas y, por tanto, cuál era el mejor momento para comprarlas.

"15 años después, podemos controlar el comportamiento de mariscos, peces y otros animales de manera más directa y precisa que a través de la temperatura", destaca Meyer en referencia a la utilidad de los sensores actuales. Entrando en la esfera humana, ya hay muestras de que se pueden predecir o inferir aspectos de la propia personalidad no revelados directamente, como nuestro [perfil psicológico](#), o [sobre las personas que nos importan](#), incluso hasta el punto de sugerirnos acciones



relacionadas con ellas, como que escribas o llames a un amigo antes del partido de fútbol de cada semana para que no se le olvide.

Estas aplicaciones hacen sus predicciones basadas en *big data* con información que ya está ahí pero también podemos realizarlas aunque no dispongamos de los datos, generando situaciones que nos los proporcionen. La **experimentación conductual** puede resultar una buena opción para obtenerlos, según el experto en análisis económico de la Universidad de Berkeley **Gregory La Blanc**.

La Blanc explica que hay muchas técnicas posibles en función de las necesidades, y pone algunos ejemplos de experimentos. En uno de una web de descargas pudieron comprobar que el botón verde tenía una tasa de conversión un 34% mayor que el botón rojo. En otro de una compañía sudafricana de crédito se dieron cuenta de que incluir una cara femenina en sus anuncios aumentaba más la conversión que una reducción del 2% de los intereses.

¿Hay transparencia en la toma de decisiones o se manipulan las opciones?, ¿qué pasa cuando los participantes saben que se trata de un experimento y responden estratégicamente?, ¿dónde queda la responsabilidad ética?

La utilidad de estos experimentos para intentar cambiar un comportamiento (botón verde para obtener más descargas) y para rentabilizar al máximo un producto (ganando un 2% más de beneficios con solo incluir una cara femenina en un anuncio) es obvia. Es lo que se llama *data-driven decision making* o toma decisiones basadas en datos. **"Cada compañía en EE.UU. y en todo el mundo está usando o usará la experimentación para medir los comportamientos de sus clientes"**, estima La Blanc.

Estamos hablando de cómo predecir y cambiar comportamientos relacionados con el consumo pero el uso de *big data* se puede aplicar a muchos otros ámbitos. Por ejemplo, al político. El uso combinado de análisis de datos masivos con la psicología conductual y el método experimental fue clave en las campañas electorales que dieron a Barack Obama a la presidencia de EE.UU. en 2008 y en 2012.

En el libro [The Victory Lab](#), el analista político Sasha Issenberg explica cómo el Partido Demócrata "sustituyó el instinto de los sabios asesores por un nuevo paradigma estratégico basado en datos". A través de entrevistas de *call centers* y de datos extraídos de los registros de inscripción de votantes, de bases de datos de consumidores y de los contactos de la última campaña electoral, lograron recopilar mil variables diferentes de dichos votantes, como sus intereses, preocupaciones o amistades o contactos. **Usando modelos estadísticos lograron no solo derivar datos a nivel individual del electorado sino también predecir su voto.**



Chris Meyer ▲
Fundador de Monitor Talent y Patrono de
Fundación Innovación Bankinter



Gregory La Blanc ▲
Investigador Haas Economic Analysis en
la Universidad de Berkeley



Antonio Damasio

Profesor y Director del Brain and Creativity Institute en la Universidad del Sur de California (USC) y Premio Príncipe de Asturias de Investigación Científica y Técnica

Todo esto tiene, por supuesto, **implicaciones éticas**. Según asegura el propio autor del libro, "realizaron experimentos aleatorios que trataban a los votantes como involuntarios conejillos de indias". Los asistentes al FTF plantean varias dudas en relación a este tipo de experimentación: ¿hay transparencia en la toma de decisiones o se manipulan las opciones?, ¿qué pasa cuando los participantes saben que se trata de un experimento y responden estratégicamente?, ¿dónde queda la responsabilidad ética?. La Blanc responde que este "juego" ya existía antes del *big data*, "con la diferencia de que ahora estas campañas pueden ser mucho más efectivas".

La publicidad es otro campo que se beneficia enormemente de esta ciencia, dando lugar a lo que muchos llaman **neuromarketing**. Se trata de una combinación entre neurociencia y minería de datos, que integra el aprendizaje computacional automático con la intuición humana y el *big data*. Esto permite una medición directa de los pensamientos del consumidor acerca de un anuncio para conocer sus efectos persuasivos, teniendo en cuenta que serán mayores cuanto más se alineen emocionalmente con el espectador. Tan solo es necesario observar si ante estos impactos se activan o no las zonas del cerebro de los consumidores que involucran las áreas de atención, emoción, memoria o significado personal.

La posibilidad de conocer -incluso en tiempo real- la respuesta afectiva a estímulos de acciones de *marketing* es un hecho. Sin embargo, la extrapolabilidad de los resultados no está tan clara. **Antonio Damasio**, fundador del Instituto del Cerebro y la Creatividad de la Universidad del Sur de California, es cauteloso al respecto. Asegura que "se

trata de una ciencia muy nueva" y que "adolecemos de un exceso de confianza en estos sistemas".

Damasio explica que "por mucho que sepamos cómo provocar una reacción y qué pasos seguir, esto solo funcionaría con seguridad si se lo aplicásemos a una máquina". "Los humanos somos seres biológicos y cada célula en nuestro cerebro es un organismo vivo y autónomo con su propio ciclo de vida, algo incontrolable a lo que añadimos un cerebro que filtra todo con un juicio personal que los algoritmos de estos sistemas no conocen ni tienen en cuenta", explica el neurocientífico.

De modo que cuando hablamos de personas no es tan fácil como decir o hacer H o B para que se produzca X o Y. "Tampoco seremos capaces de convencer a la gente de hacer lo correcto proporcionando conocimiento sobre lo que le conviene porque tenemos motivaciones, deseos y apetitos extremadamente poderosos que activan nuestro sistema de recompensa y luchan contra esa información objetiva", argumenta Damasio. El neurólogo ejemplifica así lo complejo que es cambiar comportamientos, y deja la solución en manos de su filósofo admirado, Baruch Spinoza: "La única manera realmente efectiva de combatir una emoción negativa es con otra positiva de mayor intensidad".

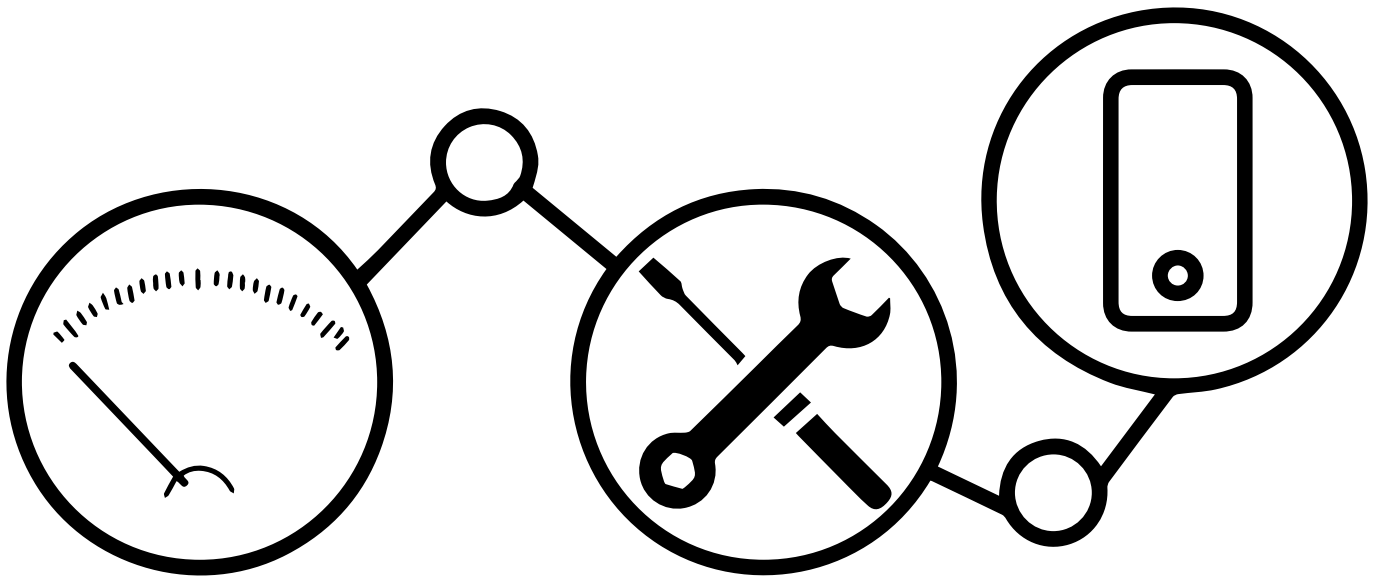
Neuromarketing.

Se trata de una combinación entre neurociencia y minería de datos, que integra el aprendizaje computacional automático con la intuición humana y el *big data*. Esto permite una medición directa de los pensamientos del consumidor acerca de un anuncio para conocer sus efectos persuasivos, teniendo en cuenta que serán mayores cuanto más se alineen emocionalmente con el espectador.

Optimizar

2.3

Pasar de indicadores clave de rendimiento (KPI) y cuadros de mando a análisis de rendimiento e ingeniería de datos.



Las predicciones del *big data* también son aplicables al trabajo interno de la organización. Los datos procedentes de internet, de medios de comunicación social y de nuestros teléfonos y dispositivos permiten medir el rendimiento con mayor precisión y usar esos conocimientos para predecir el rendimiento futuro, así como identificar y evaluar nuevas oportunidades y riesgos.

Estas aplicaciones han llegado para transformar la estructura de la organización y hacerla más ágil. Justo a tiempo para responder a una necesidad de

cambio acelerado y permanente adaptación a las nuevas -y constantes- demandas del mercado. Y justo a tiempo para sustituir a desfasados medidores y **pasar de indicadores clave de rendimiento (KPI) y cuadros de mando a análisis de rendimiento e ingeniería de datos.**

Gestionar y supervisar la ejecución de la estrategia en las organizaciones pasa ahora por una evolución de la metodología de gestión de procesos de negocio gracias a las cada vez mayores cantidades de datos a las que ahora tienen acceso. Para ello,



Taryn Sullivan ▲
CEO de Efficiency Exchange



Stefan Bungart ▲
Líder de GE Software Europe en
GE Global Research

“El objetivo es mejorar la eficiencia, reducir los costes y hacer las cosas más fáciles para la gente que trabaja en GE”.

es preciso combinar herramientas innovadoras de gestión financiera, análisis predictivo e inteligencia empresarial con análisis *big data* (predictivos y causales) en una plataforma de gestión integrada. Esta debe ser capaz, a su vez, de trasladar los resultados a informes de rendimiento interactivos y visuales, con formato digital y adaptados a los dispositivos móviles.

Se abren así nuevas oportunidades gracias a nuestra capacidad de utilizar y analizar grandes cantidades de datos (como registros web, datos de nuestros dispositivos móviles y teléfonos, estadísticas de motores de búsqueda, etc.) y fuentes de datos menos estructurados (como conversaciones en medios sociales como Facebook y Twitter, videos, etc.).

Hoy día ya hay ejemplos de empresas que integran enfoques tradicionales con análisis innovadores para analizar su rendimiento e impulsar la toma de decisiones basada en hechos. Efficiency Exchange (EEx) ayuda a muchas de ellas. Concretamente, en el campo de la manufactura. “Es algo muy poderoso y que no tiene visibilidad hoy en día; mucha gente no se da cuenta de que hay muchos más datos de fabricación que cualquier otro tipo”, asegura **Taryn Sullivan**, su fundadora y directora. Los resultados hablan por sí mismos: “Tenemos unas **300 fábricas en China ahorrando aproximadamente un 11% del coste de la electricidad**”, señala.

¿Cuál es el secreto? “Usamos los datos para crear cambio humano a nivel de fábrica, interactuando con todos los empleados en todos los contextos (incluido el tiempo de deporte y ocio) para entenderles bien y construir una interfaz sobre problemas reales que ellos tienen”, explica la fundadora de EEx. “Creo que el aspecto humano es una de las piezas fundamentales que se pierden cuando hablamos de *big data*: entender al usuario”, sostiene.

“Cuando la gente me pregunta cuál es la tecnología que nos hace tan únicos respondo que en realidad las herramientas están ahí pero no son usadas. Esto tiene que ver con la experiencia, con las

habilidades de quien está recibiendo esos datos. Por tanto donde hay que poner el foco no es en la tecnología en sí sino en la **gerencia**, en cómo usar estos datos para que la gente dirija las fábricas", comenta Sullivan.

Según esta directiva, una de las claves para mejorar operacionalmente la forma en que las fábricas trabajan pasa por entender el coste de la electricidad más allá de la energía. "Es el tipo de cosas que funcionan en todo sitios, teniendo en cuenta no solo la eficiencia energética per se sino la eficiencia de recursos, la gestión laboral y todas las utilidades. Se trata de obtener un modelo escalable, aplicable no solo a una fábrica sino a toda la cadena de producción", indica.

Entre los principales beneficios de aplicar este método Sullivan destaca la automatización de procesos, el refinamiento de la planificación y de la forma en que la fábrica opera en sí misma y la

"¡No sabemos lo que no sabemos!, Lo que sí sabemos es que este es el camino para generar ventajas competitivas verdaderamente poderosas".

reducción de los tiempos de espera y de parada. "El problema es que es solo aplicable a ciertos tipos de manufactura", asegura, "aunque entre ellos se incluyen los relacionados con la producción de la mayoría de los bienes de consumo", señala la fundadora de EEx.

General Electric (GE) es otro ejemplo de empresa que aplica *big data* a nivel operacional para mejorar rendimiento interno de la compañía. **Stefan Bungart**, director de software en la división de Investigación Global de GE, asegura que para ellos son prioritarios estos proyectos que "generan ahorro interno de costes y mejoras para las empresas de servicios (eficiencia y efectividad de la fuerza de ventas, perfeccionamiento del diseño de ingeniería, monitoreo de equipos a distancia y diagnósticos), sin tiempo de inactividad no planificado". "El objetivo es mejorar la eficiencia, reducir los costes y hacer las cosas más fáciles para la gente que trabaja en GE", indica.

Bungart busca en estos proyectos "una vía optimizada de hacer llegar la eficiencia a todos los rincones de la empresa, de forma envolvente". El proceso a seguir pasa por obtener datos de valor sobre fallos o potenciales errores operativos, extrapolarlos al futuro para predecir qué pasaría de seguir así, y mandar a los técnicos a repararlo antes de que se produzca el problema. Una especie de uso preventivo del *big data* para optimizar el servicio de operaciones, mejorar la productividad y ayudar a los técnicos a ser más efectivos.

Según explica el directivo de GE, la manera en que lo están llevando a cabo es mediante una combinación de aprendizaje computacional automático e interacción humana, de modo que un algoritmo que ha sido creado por una persona se mejore a sí mismo constantemente a través de esta reciprocidad.

Bungart también apuesta por el análisis exploratorio de datos: "¡No sabemos lo que no sabemos!", enfatiza. Lo que sí sabemos es que este es el camino para **generar ventajas competitivas verdaderamente poderosas**.

Innovar para transformar

2.4



Lita Sands

Líder global de Transformación Digital de Novartis

El *big data* puede ayudar a GE a optimizar sus procesos internos pero también puede ofrecerles una nueva manera de aproximarse al consumidor, de establecer otro tipo de relación que les aporte valor más allá de la oferta de un producto concreto. Bungart sabe que esta es la llave para pasar de la tradicional innovación iterativa a una innovación disruptiva. Y el primer cambio es sin duda la apertura y el cambio de mentalidad corporativa. Apertura para abrir los problemas de la organización y buscar soluciones colectivas, usando la inteligencia de la masa mediante el *crowdsourcing* y la innovación abierta. Y cambio de mentalidad corporativa para pasar de dirigirse al cliente a dirigirse a las personas, a los ciudadanos.

Eso es exactamente lo que han hecho en GE en varios de sus proyectos más recientes. Por ejemplo, en un trabajo conjunto con Alaska Airlines que se proponía economizar el gasto de combustible de la aerolínea. "Cuando comenzamos a observar los datos contextuales para encontrar una forma de reducir el retraso global encontramos una deficiencia en nuestro plan de hacer esto solos; no sabíamos por dónde empezar ni qué buscar", relata Bungart. "Así que decidimos recopilar todos los datos que podíamos, subirlos a internet y **abrir el problema a todos los científicos de datos del mundo**, de modo que cualquiera pudiera analizar y crear su propio escenario", explica. A cambio, la persona o el equipo que mejorase su algoritmo para hacer unas predicciones fiables y optimizar el problema de modo escalable, obtendría un premio de 100.000 dólares.

El resultado: los ganadores del [GE Flight Quest](#) lograron reducir 2,5 minutos de tiempo en cada vuelo, lo que supone un **ahorro anual de combustible y personal de 26 millones de dólares** debido a la liberación de dos aeronaves en una compañía con una flota media de 120 aviones. 200 equipos de todo el mundo participaron en esta competición lanzada a través de [Kaggle](#), la comunidad más grande de científicos de datos enfocada en la resolución colectiva de problemas.

Junto a este, GE ha lanzado otros proyectos que han generado una reducción de costes y un aumento de productividad, además de aportan valor a los usuarios finales. Por ejemplo en el ámbito de la integridad de los ductos de petróleo y gas, en el del rendimiento de la red ferroviaria y también en el de la gestión de la dosis de radiación en el sector

salud. Para Bungart, el éxito de estas iniciativas no solo reside en la rentabilidad obtenida sino en la implicación y la pasión de los participantes. "Puedes ver el fuego en sus ojos, el entusiasmo de participar en algo grande", enfatiza el directivo de GE.

Un verdadero cambio en la cultura de la empresa es el que ha experimentado la farmacéutica Novartis. **Lita Sands**, directora de Transformación Digital, cuenta su caso: "Quiero hablaros del camino que llevamos recorriendo los últimos tres años. En solo tres meses desde que comenzamos este proceso eliminamos por completo el papel (y lo sustituimos por portátiles y por tabletas para la fuerza de ventas) e integramos los sistemas de venta en formato digital. Fue algo emocionante y aterrador al mismo tiempo", reconoce Sands.

Habían pasado
a la acción pensando
en ese por qué pero
sin tener en cuenta
el cómo.

¿Cómo lo hicieron? "Al estilo Apple. Nos dimos cuenta de que había que hacer algo sencillo para que cualquiera pueda adoptar el cambio. Así que usamos la inspiración de los periódicos -algo tan simple como un sitio donde puedes leer historias- para nuestra plataforma de lanzamiento. Cuando compras un coche no tienes que aprender a conducir de nuevo. Esa era nuestra filosofía", responde Sands.

"De un día para otro pasamos de tener a toda la fuerza de ventas trabajando con papel a contar con un sistema en el que, gracias a la digitalización, podemos controlar todo lo que compartimos, en qué están interesados los doctores, cuáles deben ser las próximas transacciones y qué deberíamos estar haciendo como marca comercializadora", cuenta la directiva. "Sin embargo -señala- todo el mundo parecía estar encantado al principio pero cuando observamos los datos reales de cada país y de cada región vimos que nada de eso se reflejaba en ellos".

"Nos dimos cuenta de que solo estábamos atajando el *por qué*, porque queríamos una transformación digital", asume Sands. No era suficiente solo con darles unas tabletas a los empleados y pensar que con eso estaba todo hecho. **Habían pasado a la acción pensando en ese por qué pero sin tener en cuenta el cómo:** "El cambio digital lo es todo. La gente se da cuenta de que es una necesidad y que tenemos que hacer un esfuerzo por abrir un diálogo real para pasar del papel al multimedia y que las cosas sigan funcionando". Con esta premisa, buscaron una forma de abordar el cambio, combinando esos *por qué* y *cómo* a la vez.

El modelo teórico a seguir fue la gestión del cambio en ocho fases de John Kotter, profesor de la Escuela de Negocios de Harvard. Este implica establecer un sentido de urgencia (necesidad de abordar dicho cambio), formar un equipo que lo lidere, desarrollar una visión que lo dirija, comunicarla y transmitir la estrategia a seguir, capacitar a los demás para actuar en consecuencia, asegurar resultados a corto plazo, consolidar y profundizar en las mejoras e institucionalizar los nuevos métodos para un continuo desarrollo.

Decidimos recopilar todos los datos que podíamos, subirlos a internet y abrir el problema a todos los científicos de datos del mundo.

Gestión del cambio en ocho fases de John Kotter, profesor de la Escuela de Negocios de Harvard.

Este implica establecer un sentido de urgencia (necesidad de abordar dicho cambio), formar un equipo que lo lidere, desarrollar una visión que lo dirija, comunicarla y transmitir la estrategia a seguir, capacitar a los demás para actuar en consecuencia, asegurar resultados a corto plazo, consolidar y profundizar en las mejoras e institucionalizar los nuevos métodos para un continuo desarrollo.

El ejemplo práctico que les inspiró fue el caso de Burberry, cuya transición digital lideró su directora general, Angela Ahrendts, ahora vicepresidenta de venta al por menor y *online* de Apple. "Ahrendts se dio cuenta de la importancia de usar los medios sociales y de contar con una buena plataforma de relación con el cliente (CRM) para digitalizar sus tiendas minoristas", cuenta Sands. "Renovó la cultura de los empleados, creó todo tipo de comités para ayudar a entender el cambio y transferir sus habilidades laborales hacia las nuevas necesidades de creación y gestión de contenidos, de interacción social media y de creación de comunidad, así como del *back-office*", cuenta Sands.

Otro caso inspirador fue el de la cadena de supermercados Tesco. En su afán por lograr liderazgo en Corea del Sur sin construir nuevas tiendas descubrieron que uno de los principales problemas de la gente allí es que no tiene tiempo para comprar. Así que **decidieron llevar las tiendas a las personas mediante un sistema de compra virtual en el metro**. El sistema -bautizado como Homeplus- recrea los estantes del supermercado en las paredes de los pasos subterráneos, de manera que cualquier persona que pasa por allí puede usar su smartphone para seleccionar los productos allí reflejados. Tras realizar la compra *online*, el supermercado se encarga de llevarle los productos a su casa. El éxito de Homeplus dio a Tesco el liderazgo entre los supermercados *online* en Corea del Sur y el segundo puesto en el ranking *offline*.

Con la base teórica de Kotter y los ejemplos prácticos de Burberry y Tesco, Novartis reinició su transición digital, pasando del *por qué* al *cómo*. "Creamos una un comité operativo y una fuerte estructura de gobernancia y trabajamos con cada región para centrarnos en sus principales preocupaciones", detalla Sands. "Nos pusimos en el lugar de la fuerza de ventas para entender sus frustraciones y nos dimos cuenta de que el acceso a los datos y su análisis era siempre problemático", prosigue. Así que crearon unos cuadros de mando "muy sencillos" y lo primero que hicieron fue explicar a los empleados

como usarlos. "En Italia su adopción en un mes pasó del 1% al 30% de los empleados, y lo mismo sucedió en China y en otros países", resalta Sands.

"Nos dimos cuenta de que hay comportamientos humanos muy simples que no estábamos apreciando y nos centramos en ellos para hacer el la transición lo más fácil posible", reflexiona la directiva de Novartis. "Identificamos dónde estaban las competencias, cuáles eran únicas y útiles para el futuro, y lo que encontramos básicamente fueron dos cosas: curiosidad y pasión. Y trabajando con la gente y entrenándola es cómo lo digital se está convirtiendo en el ADN de nuestra compañía", afirma Sands.

La transformación con médicos, pacientes y empleados continuará para seguir mejorando esta maravillosa experiencia de usuario. Seguiremos preguntando a la mente y tocando el corazón de los empleados para que se suban al barco.

"De esta manera, hemos adoptado un nuevo sistema con el que podemos controlar cada vez que alguien hace clic en algo y también la respuesta del doctor, con lo que estamos creando un indicador que correlaciona el sentimiento del cliente con el comportamiento de prescripción", detalla Sands. "Gracias a ello hemos comprobado cómo cada 1% de aumento en satisfacción del cliente se tradujo en un aumento de ventas del 6,5%, lo cual es muy potente", añade. Las analíticas en tiempo real les han permitido **mejorar la experiencia de usuario y obtener información para seguir optimizándola**. "Este conocimiento nos permite tomar decisiones basadas en datos y pasar a la acción de forma eficiente", afirma Sands.

"Cuando haces todas estas cosas juntas puedes hacer crecer tu organización más rápido con oportunidades exponenciales facilitadas por la digitalización, que ya es un área fuerte de la compañía", asegura Sands. "En tres años hemos sido capaces de transformar la organización de modo que 50.000 individuos están realmente usando sus capacidades digitales, siendo proactivos y queriendo innovar", destaca la directiva de Novartis. "La transformación con médicos, pacientes y empleados continuará para seguir mejorando esta maravillosa experiencia de usuario. Seguiremos preguntando a la mente y tocando el corazón de los empleados para que se suban al barco", concluye Sands.

Así que cuantificar datos para obtener información valiosa para el negocio, predecir comportamientos y cambiarlos, optimizar el rendimiento operativo de las organizaciones e innovar para una transformación que lleve a una verdadera ventaja competitiva -y a un aumento de los beneficios- son solo algunos de los superpoderes que ya está ofreciendo el *big data*. Sin embargo, el camino apenas acaba de comenzar y **quedan por delante muchos retos que afrontar y obstáculos que superar**, como veremos en posteriores capítulos.

Big data, big money. O eso dicen.





De personas para personas

3

Cap. 3/7

- 3.1 Gobernantes y ciudadanos
 - 3.2 Médicos y pacientes
 - 3.3 Directivos y empleados
 - 3.4 Entrenadores y jugadores
 - 3.5 Policías versus criminales
-



De personas para personas

Vicki Seyfert-Margolis. Fundadora y CEO de My Own Med, Inc.

➔ **Casi sin saberlo,** venimos participando como sociedad en una revolución cultural que se inició en internet y que han abanderado empresas como Amazon, Google y Facebook—empresas que son ya una parte natural de las vidas de muchos de nosotros.

El contrato social que hemos firmado gira en torno a la comodidad de tener acceso instantáneo a información, bienes, personas y conectividad a cambio de información sobre nosotros. Es cierto que el contrato desde nuestro punto de vista, como usuarios finales, es extremadamente valioso. No obstante, el valor para las empresas es aún mayor. Nuestros datos se albergan en grandes servidores centralizados, convirtiéndose así en el activo sobre

el que reposa el extraordinario valor de estas empresas. Por tanto, la pregunta es: ¿el contrato social es igualitario y los datos de las personas—la razón de ser del valor que ostenta la minoría que alberga los datos—se intercambian justamente por servicios para las personas?

Podría razonarse que esta cuestión ha ido evolucionando con el paso del tiempo: el valor de

¿El contrato social es igualitario y los datos de las personas se intercambian justamente por servicios para las personas?

las empresas se disparó según se fue extendiendo el uso de las herramientas. Y esto ocurrió de forma casi imperceptible para cada persona que contribuía con sus datos. En el momento actual ponemos en duda quién tiene derecho de poseer y usar nuestros datos, y si lo que nos ofrecen a cambio de nuestra contribución es suficiente. Conforme big data empieza a aplicarse en sanidad (nuestros datos más personales, podría decirse) la cuestión se está planteando como un tema de privacidad. Es irónico que a través de nuestras compras online, participación en redes sociales y la economía de las aplicaciones ya hayamos cedido una cantidad ingente de datos personales. No obstante, la salud es el catalizador que pone en duda el statu quo.

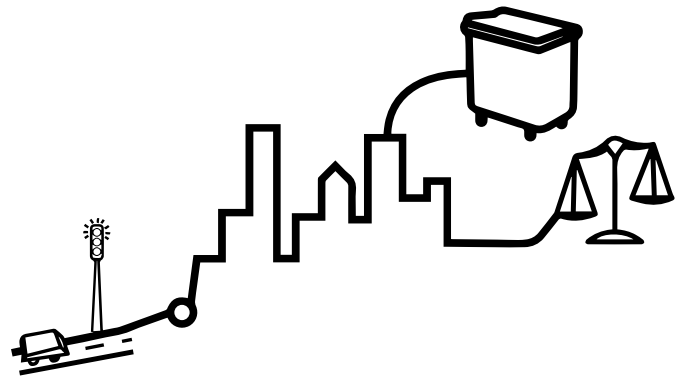
Dado el uso creciente de los sensores en dispositivos wearable (que vestimos), los historiales médicos electrónicos, las aplicaciones relativas a la salud y al bienestar, y hasta la secuenciación del genoma humano; nuestro "yo" más profundo, aquello que nos hace lo que somos, se puede capturar en tiempo real a través de un dispositivo tan sencillo como un reloj o una pulsera, o tan sofisticado como los monitores de glucosa y las bombas de insulina que han automatizado la gestión de la diabetes. Gracias a los datos se sabe dónde vamos, cuánto nos movemos, cuánto dormimos, nuestra frecuencia cardíaca y mucho más, y todo ello se viene combinando con la información de nuestros historiales médicos para definir nuestro estado de salud actual, predecir nuestra trayectoria de salud y evaluar los riesgos. Visto así, es el big data de la persona, para la persona en su mejor versión, puesto que podemos empezar a soñar con intervenciones de salud en tiempo real y medicina preventiva en contextos hasta ahora inimaginables.

¿Y qué pasa con el contrato social? La pregunta ante nosotros como individuos y como grupo se reducirá a evaluar el riesgo y el beneficio de las decisiones más personales sobre salud, tal y como lo definan los datos. Al plantearnos la vida con big data y sensores que calculan nuestro estado de salud, es obligatorio preguntarse: ¿cuántos datos más queremos ceder y a cambio de qué?

Introducción

3.0

➔ **Hablar únicamente** de las aplicaciones del análisis masivo de datos en los negocios sería muy reduccionista. Al fin y al cabo, lo que de verdad importa es cómo esto sirve a las personas. No en vano, cada uno de nosotros somos una fuente de datos, una de las más relevantes, de hecho, de entre las que se miden y se ponen a disposición de las herramientas de análisis.



Las posibilidades de estos datos que se extraen "de las personas para las propias personas" y para la relación entre ellas son múltiples, tanto en el ámbito de la salud como en el del gobierno o en el de los recursos humanos, para optimizar y potenciar nuestro rendimiento y para construir unas ciudades más seguras y habitables. Son varios los escenarios

de aplicación del *big data* 'de personas para personas' y para la relación entre ellos, ya sean gobernantes y ciudadanos, médicos y pacientes, directivos y empleados, entrenadores y jugadores o policías *versus* criminales. Sobre cada uno de ellos se abre un grupo de trabajo en el FTF, cuyas aportaciones reflejamos y ampliamos a continuación.

Gobernantes y ciudadanos

3.1

Tal y como exponen los expertos reunidos en este grupo, las nuevas tecnologías basadas en recolección y procesamiento de datos confieren a los políticos y a los gestores de las ciudades de una capacidad de escucha, acceso y conocimiento de sus problemas y necesidades antes impensable. Y no solo eso sino también de proporcionar una respuesta. Permiten un **flujo de comunicación de abajo hacia arriba y de arriba hacia abajo**, de modo que los ciudadanos puedan participar en una toma de decisiones colectiva. Esto es: el sueño de la democracia directa, hasta hace poco considerada una utopía.

"Esperamos que muchos de los nuevos datos recogidos hablen por sí mismos y que las ciudades respondan a ellos", aseguran desde el FTF. Tal y como señalan, la recopilación pasiva de datos que indican cosas como que hay problemas en el parque local va a posibilitar un gobierno más adulto, más receptivo a las necesidades y a darles respuesta. Sin embargo, también se plantean dudas sobre la priorización de los políticos ante estos problemas y si se centrarán en aquellos de más fácil respuesta a corto plazo o en los que realmente aportan valor a largo plazo.

"La tecnología está acelerando la capacidad de escuchar los problemas. A cambio, se generan expectativas de respuesta rápida por parte de los poderes públicos", señalan los expertos. Esta presión puede llevar a tomar decisiones precipitadas para resolver problemas tal vez secundarios o menos relevantes que otros que requerirían más tiempo. Además, a menudo se proponen por parte de la ciudadanía soluciones inviables o no óptimas por falta de conocimiento acerca de los recursos necesarios y de otras posibles alternativas, o

La tecnología está acelerando la capacidad de escuchar los problemas.

simplemente porque son las más fáciles y rápidas de implementar para solucionar en apariencia el asunto. Sin embargo, sería un error que los gobernantes cedieran a estas presiones cuya respuesta es solo un parche y no una solución a largo plazo.

El reto en este sentido -apuntan en el FTF- está en analizar cómo dar a la gente una retroalimentación de valor sobre estas discusiones de alcance y que puedan tomar decisiones más informadas en cuanto a su opinión sobre ellas y a cómo se debería actuar.

Un aspecto clave en esto es el de la transparencia.

Los presentes coinciden en que una política de datos abiertos con un gobierno abierto permite que los ciudadanos tengan acceso a la información y por tanto es una herramienta fundamental para la toma de decisiones. "Además es un arma contra los rumores que destruyen la reputación de mucha gente, y esa es una utilidad en la que a menudo la gente no piensa", añade una de los participantes en alusión a las reticencias de determinados estamentos públicos a abrir sus datos.

Entonces, ¿cómo reunir todos estos datos y crear medios para que los ciudadanos expresen sus preocupaciones sobre la ciudad? Algunos en el grupo de trabajo proponen crear algo parecido a un mercado *online* como Amazon, un equivalente para la gestión gubernativa fácil de entender y de usar. "Actualmente cuando tenemos un problema escribimos al alcalde para quejarnos y no se entera nadie más. En Amazon cuando se escribe un comentario sobre un producto otras personas lo ven y pueden sentirse identificadas y sumarse", resalta uno de los expertos. "Además, la compañía es especialmente sensible a estas revisiones públicas, que suelen obtener una rápida respuesta en términos de mejora del producto o del sistema de envío, dependiendo de cuál fuera el problema señalado", destaca.

El debate a la hora de crear una plataforma de este tipo surge cuando se propone incorporar un *rating* -un sistema de valoración- aplicado a cuestiones políticas en el que los ciudadanos otorguen a los diferentes temas un número determinado de estrellas. "Dar estrellas a cuestiones políticas puede no ser una buena solución", sostiene uno de los participantes. No obstante, en general hay acuerdo sobre la **necesidad de este tipo de espacios para la participación ciudadana.**

En realidad ya existen numerosas iniciativas de este tipo. Algunos ejemplos son las plataformas de gobierno abierto de [Canadá](#) o de [Estados Unidos](#). En Europa, estos programas a menudo se canalizan a través de portales de datos abiertos, que en realidad son solo una parte de lo que se considera

el gobierno abierto. No obstante, si hay numerosos proyectos regionales y municipales en este marco de gobernanza. En España, algunos ejemplos son el portal [Irekia](#), del País Vasco, o el de la provincia de [Zaragoza](#). En el ámbito de los datos abiertos se creó en marzo de 2010 el portal [datos.gob.es](#), aunque a nivel autonómico y local hay multitud de proyectos.

Estas plataformas, sin embargo, tienen una interfaz a menudo poco amigable para el usuario, con demasiado texto y dificultad para acceder a lo que se busca, lejos del estilo Amazon. En el grupo de trabajo del FTF creen que todavía **son necesarios más esfuerzos en torno a la transparencia y el empoderamiento ciudadano.** "La alternativa es seguir como hasta ahora, manifestándonos únicamente con un voto una vez cada 3 o 4 años. ¿Es ese el tipo de dinámica democrática que queremos?, cierra a modo retórico uno de los presentes.

Necesidad de este tipo de espacios para la participación ciudadana.

Médicos y pacientes

3.2

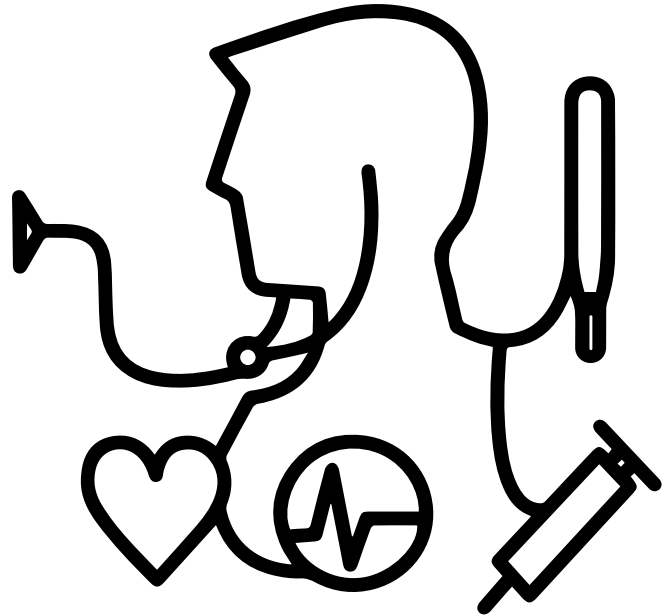


Vicki Seyfert-Margolis
Fundadora y CEO de My Own Med, Inc.

El *big data* está contribuyendo a reinventar el sistema sanitario.

El *big data* está contribuyendo a reinventar el sistema sanitario. Es lo primero que destacan los expertos del FTF en el grupo de trabajo sobre la influencia de la medición y análisis de datos masivos en la salud y en la relación entre profesionales sanitarios y pacientes. "Parte de la visión de cómo son las organizaciones de salud y cómo se proveen los cuidados sanitarios está en un proceso de grandes cambios. Se están produciendo en los negocios, en la atención a los pacientes y en la industria en sí, con grandes disrupciones", señalan.

"Hay innovación a varios niveles, que influye en la práctica de los hospitales, de las enfermeras, de los investigadores... Una de las cosas que es diferente en relación con los datos es que hay una transformación del tradicional seguimiento periódico del enfermo en consulta a un proceso continuo en el que, **a través de una plataforma digital o de una app, los**



profesionales pueden controlar minuto a minuto la evolución del paciente y realizar mediciones constantes. Se produce un flujo continuo de procesamiento de datos, y en cualquier industria que pasa esto todos los negocios cambian", apuntan.

Los sensores incorporados en nuestros *smartphones*, en pulseras y relojes o en otros *wearables* proporcionan un *feedback* biológico y pueden ayudar tanto al seguimiento y mejora de patologías concretas como de la salud integral del paciente. Pero según **Joseph Kvedar** para que el uso de estas aplicaciones funcione deben cumplirse tres condiciones: predisposición al cambio, adopción de herramientas de datos de salud conectadas y práctica. Este es, de hecho, el objetivo que persigue el centro que dirige en la Escuela Médica de Harvard: promover entre las personas una actitud abierta hacia las nuevas tecnologías de salud personal.

“Lo importante no es mirar a los datos sino relacionarlos con el contexto y la situación del paciente y cómo está actuando”.

Tanto para profesionales sanitarios como para pacientes la integración de estas tecnologías supone un cambio de cultura y de comportamiento y también del rol que juegan en la atención. **No solo se ve afectada la relación entre médico y paciente sino dentro de la familia y de la comunidad.** “El papel de las comunidades -tanto físicas como en internet- es un factor de cambio decisivo porque gracias a los datos pueden darse cuenta de que algo falla y pueden actuar para modificar comportamientos poco saludables de sus integrantes”, señalan en el grupo del FTF.

“Esto tiene que ver con los roles que la gente percibe que son importantes en la comunidad. Aquellos que los ejercen son escuchados por sus pares como lo harían con la enfermera o con el doctor. Por eso tienen el deber de promover actitudes que mejoren la calidad de vida del grupo, cuya identidad cultural ejerce a menudo más presión que las autoridades”, añade uno de los expertos.

Vicki Seyfert-Margolis, exasesora de Obama en el Consejo de Ciencia y Tecnología, coincide en que la comunidad tiene un gran impacto en cómo finalmente actuamos con respecto a nuestra salud. Su experiencia en este campo la llevó a crear [My Own Med](#): una plataforma en la que paciente, cuidador y profesionales sanitarios comparten una misma interfaz, tienen acceso a los mismos datos y pueden trabajar en la misma página usando una arquitectura en nube. “Su objetivo principal es empoderar a la comunidad y a los profesionales de la salud, conectándolos con su población en un ecosistema donde la transparencia y la confianza son esenciales”, señala su fundadora.

“Tecnológicamente, My Own Med combina sensores, *big data* y todas las herramientas al alcance para

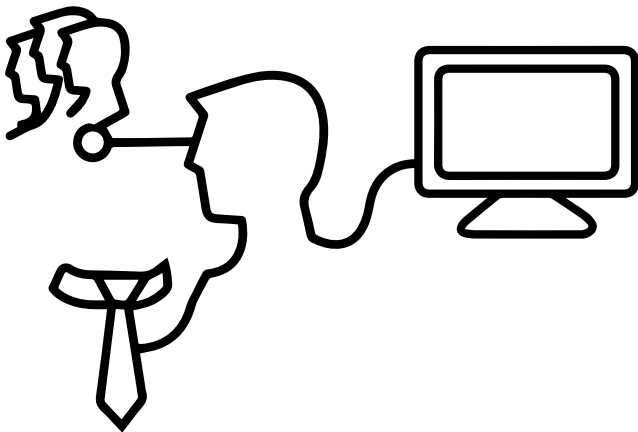
tener una visión integral del paciente y una gestión en equipo en torno a él. El aprendizaje de otros esfuerzos en *big data* demuestra que tenemos que usar estas herramientas no solo para ver los datos objetivos sino para **aprender de los sentimientos** y del estado en que se encuentran las personas a todos los niveles. Eso nos permite proveer información microsegmentada que mueva a la gente a actuar”, detalla Seyfert-Margolis.

En opinión de la experta, “lo importante no es mirar a los datos sino relacionarlos con el contexto y la situación del paciente y cómo está actuando”. “Esta integración es crítica para la próxima generación de soluciones de salud. Hay muchas posibilidades y por el momento estamos solo arañando la superficie”, puntualiza.

Directivos y empleados

3.3

Una nueva forma de organización y de relaciones entre directivos, gestores y empleados.



Stefan Bungart lo tiene claro: "No es dinero lo que motiva a los empleados". "Es relevancia, valor, sentido real, que puedan sentirse parte de un proyecto que cambiará la forma en que el mundo funciona. Esto es lo que realmente anima la gente a dar lo mejor de sí mismos", enfatiza el directivo de General Electric.

Sin embargo, las estructuras empresariales tradicionales, altamente jerarquizadas, no dan respuesta a estos anhelos y se están quedando obsoletas. Tal y como constatan los expertos del FTF, actualmente estamos inmersos en un periodo de transición hacia **una nueva forma de organización y de relaciones entre directivos, gestores y empleados**. El *big data* puede hacer este cambio más eficiente, por ejemplo a través de la implantación de una política de datos abiertos.

La información es poder y cuanto más poder tenga cada uno de los individuos y de los equipos que conforman la empresa más capaces serán de tomar decisiones por sí mismos, si es esto realmente lo que se quiere. Además, la autonomía otorga independencia y esta conlleva una responsabilidad mayor sobre las propias acciones. Esta autonomía no significa soledad: se trata de trabajar de forma individual pero en equipo, con una estructura multitarea basada en la transparencia del trabajo y de los procesos.

"¿Cómo mejora esto el rendimiento de la organización?", se pregunta **Julia Li**, consejera del laboratorio Media Lab del Instituto Tecnológico de Massachusetts (MIT) y fundadora de HCD Global, una empresa que implementa programas de

aprendizaje experimental a través de la gamificación y de tecnologías digitales. "Antes la gobernanza modelaba las compañías pero el escenario va a cambiar, empezando por el liderazgo, que pasará a ser colectivo", reflexiona Li. En su opinión, adaptarse requiere "apertura de mentes, visión de presente y de futuro y moverse rápido, ya que los datos hacen los procesos más eficientes".

Li está convencida de que **hay que destruir las estructuras basadas en el poder individual y jerarquizado**: "Es necesario para crear un sentido de urgencia para que la gente sacrifique su propia independencia y aplique su *saber hacer* y su responsabilidad en el cambio. Fomentar la apertura de mente y la colaboración". Y, sobre todo, la transparencia: "Cuando la información fluye y son públicas cosas como los presupuestos, el rendimiento y las actividades de la compañía todo el mundo puede opinar y participar. "Tenemos que pensar que contamos con personas preparadas y experimentadas que son capaces de sacar adelante la compañía por sí mismas", señala.

Maarten den Braber hace hincapié en este concepto de autonomía y responsabilidad: "Tiene que ver con una actitud muy diferente sobre cómo concibes el liderazgo y la gerencia. El modelo de autoridad del CEO representa a la compañía. Su rol es transmitir a los empleados que realmente pueden hacer las cosas por sí mismos y que no tienen que pedir permiso". Otro de los presentes refuerza esta visión del **líder como facilitador y no como autoridad**, y compara su figura con la de un director de orquesta "que transmite a los músicos dónde se dirigen y extrae de cada uno de ellos su mejor actuación".

Sin embargo, democratizar la toma de decisiones no es tan sencillo. "Incluso para empresas con una avanzada visión de futuro es un gran paso adoptar este tipo de procesos dirigidos por datos abiertos porque para facilitar esta apertura de quién toma las decisiones y cuáles son sus efectos requiere un cambio de cultura corporativa", sostiene Den Braber. Es difícil, pero no imposible. Zappos, la tienda de zapatos líder *online*, es un ejemplo de ello. La



Julia Li
Fundadora y CEO de HCD Global

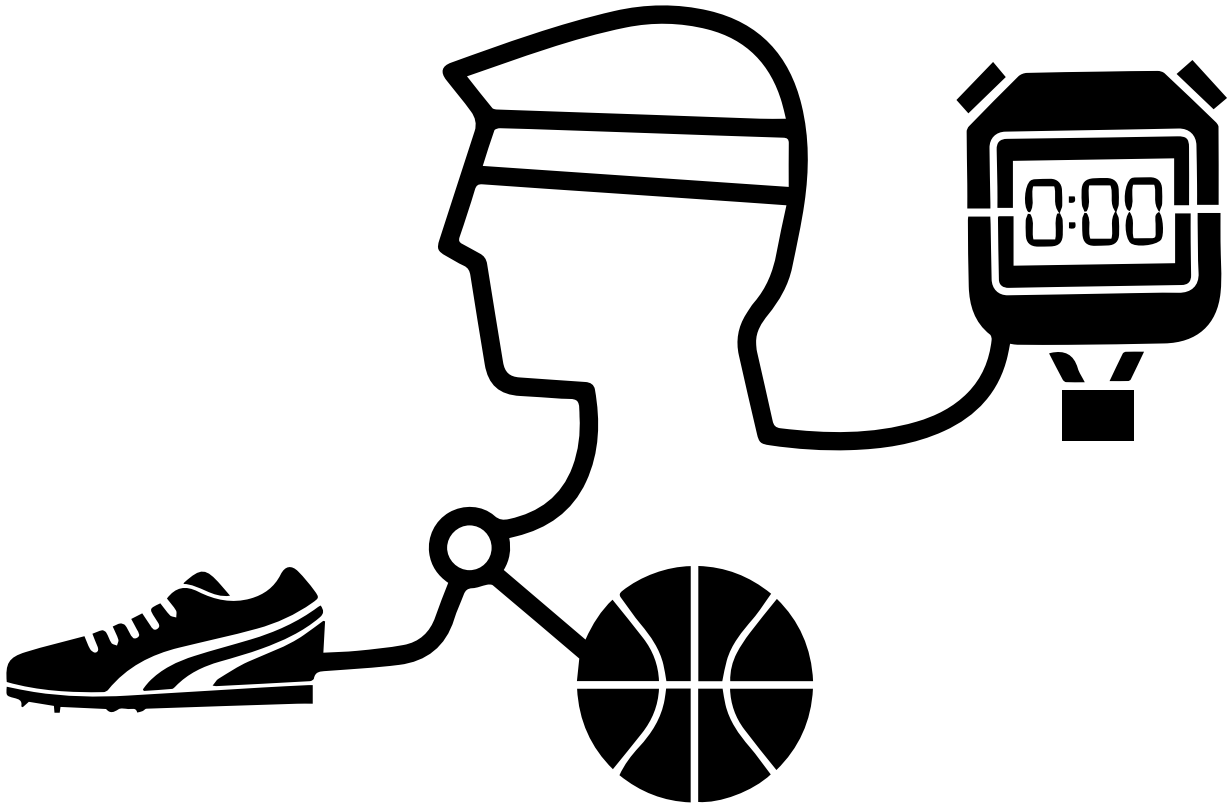
compañía se ha convertido en un ejemplo a seguir desde que a finales de 2013 anunciara su transición hacia un modelo de organización holocrático. Este modelo -la holocracia- distribuye la autoridad y la toma de decisiones en equipos autoorganizados que funcionan de forma autónoma aunque dependiente de la organización.

Este sistema no solo hace felices a los empleados sino que en combinación con el acceso a grandes cantidades de datos abiertos fomenta el **intraemprendimiento y la transformación basada en innovación**, tanto en procesos internos como de puertas afuera. Además, es una valiosa herramienta de conocimiento y observación de los competidores y del camino hacia donde avanza el sector, lo cual también genera ideas para nuevos desarrollos de negocio.

Sin embargo, hay a quienes aun confiando en el valor del *big data* no les convence la idea de horizontalidad. Algunos expertos presentes en el FTF ponen de relieve que "la democratización puede ralentizar la toma de decisiones" y actuar así como contrapeso de la optimización que ofrece el *big data*. "Debe haber un balance, una manera de usar los datos para reforzar la toma de decisiones racional, pero sigue siendo tarea del CEO. Este debe contar con la información suficiente para saber cuál es el siguiente paso porque al final es quien tiene que responder. Además, los inversores no van a dirigirse al equipo a la hora de pedir explicaciones", señalan como cierre del debate.

Entrenadores y jugadores

3.4



¿Se imaginan a un ingeniero de telecomunicaciones como entrenador de uno de los mejores equipos de fútbol del mundo? La posibilidad no es remota, y la razón es el *big data*. Analizar patrones de juego o conocer las jugadas más óptimas son algunas de sus utilidades aplicadas al deporte. Y son la razón de que **Pablo Rodríguez**, director de Telefónica I+D, esté inmerso en un proyecto de datos pionero en el mundo futbolístico en colaboración con el Fútbol Club Barcelona.

Según cuenta Rodríguez, los directivos del Barcelona se dieron cuenta de que lo importante en el campo de

juego no es mantener la estructura y el rol de cada jugador, sino conservar la pelota cuando se tiene y recuperarla cuando no. Y la ciencia computacional podía ayudar en esto desde una concepción del juego no como un conjunto de individuos sino como una red cuyo desempeño es óptimo cuanto mejor es la conexión (los pases) de los puntos que la componen (los jugadores). Por tanto, se podrían definir **nuevas estrategias para el campo de juego combinando el análisis de datos y la teoría de redes**.

"No se trata solo de saber dónde están o deben estar los jugadores sino dónde pueden estar. Hasta el

Pueden mejorar gracias al *big data* y al modelaje computacional a partir de estadísticas. Esto permite generar visualizaciones de movimiento, zonas calientes, etc., de modo que puedan anticiparse en la estrategia

momento los entrenadores se estaban guiando por la intuición y por la experimentación pero ahora lo pueden mejorar gracias al *big data* y al modelaje computacional a partir de estadísticas. Esto permite generar visualizaciones de movimiento, zonas calientes, etc., de modo que puedan anticiparse en la estrategia para ser capaces de reaccionar durante el juego", explica Rodríguez.

El interés para mejorar las estrategias de los entrenadores y para los propios deportistas es obvio. Sin embargo, ambos tienen **reticencias a incorporar tecnologías de análisis de datos** en la toma de decisiones táctica. Sobre todo los primeros, que pueden sentirse de alguna manera amenazados por una tecnología que realiza parte de su trabajo y que puede ser vista como sustitutiva en lugar de como complementaria. Por otra parte, si no hay buenos científicos y analistas de datos que sepan entrenar en el uso y aplicación de estas herramientas, se verán como algo inútil en la práctica.



Pablo Rodríguez ▲
Investigador y Director de Innovación en Telefónica



Wilfried Vanhonacker ▲
Profesor en Olayan School of Business, AUB & Patrono de Fundación Innovación Bankinter

"Es una cuestión de educación en ambos casos, tanto a los analistas de datos como a los entrenadores y jugadores. Tiene que haber una relación democrática pero esto es difícil en una estructura tradicionalmente vertical. La democratización requiere un cierto cambio de rol de los entrenadores y que participen en la limpieza de los datos, lo cual supone el 99% del trabajo. Y por supuesto son una pieza clave para asegurar que el valor obtenido está siendo transmitido a los jugadores" comenta **Wilfried Vanhonacker**, profesor de Marketing de Coca Cola y decano de la escuela de negocios OSB en la Universidad Americana de Beirut.

"Tienen que abrir su mente a los datos y a su análisis, y para ello hay que explicárselo todo: qué son los datos, cómo construyes el modelo... solo así se darán cuenta de que colectivamente pueden alcanzar mejores resultados" señala Vanhonacker por su experiencia trabajando con equipos de fútbol. Desde su punto de vista, el foco no debe ponerse en ver qué podría funcionar: "Eso podemos saberlo fácilmente. Podríamos identificar cualquier mejora en el rendimiento de los jugadores, el problema es que también pueden hacerlo los entrenadores de los equipos contrarios y averiguar así sus estrategias de juego".

Es decir, que **las jugadas se pueden volver muy predecibles**. "Así que hasta cierto punto tenemos que volver atrás y pesar cómo estratégicamente podemos reintroducir antiguas tácticas". "En definitiva, diseñar nuevas formas de entrenar entendiendo si las diferentes decisiones que tomamos en el pasado fueron útiles para ganar el juego o no", concluye.

Policías versus criminales

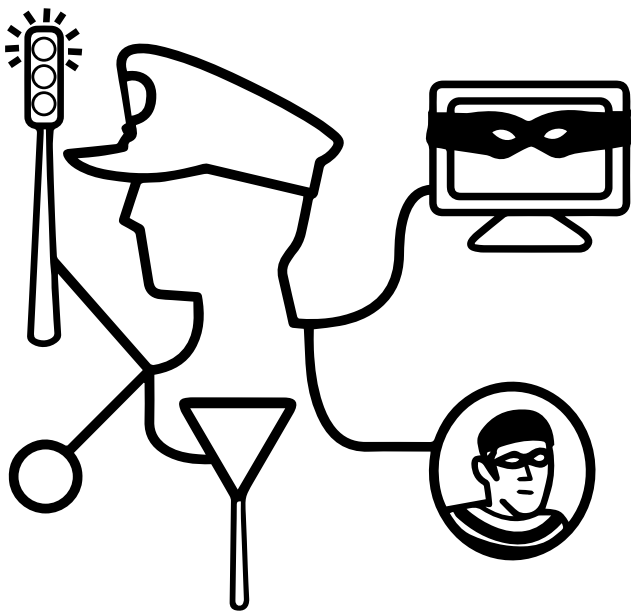
3.5

La vulnerabilidad aumenta en el caso de los dispositivos móviles.

La tecnología reproduce el crimen. Lo dice Frank Abagnale, el impostor más famoso de la historia. Este estafador al que Steven Spielberg dedicó una película (*Atrápame si puedes*, protagonizada por Leonardo Di Caprio) trabaja hoy con el FBI y con empresas privadas para combatir el fraude *online*. "Si no puedes con el enemigo, únete a él", dice la famosa frase del célebre libro de Sun Tzu *El arte de la guerra*. Y es que esa guerra hoy se libra en internet. Gracias a las nuevas tecnologías de la información, el robo de identidad es miles de veces más fácil de realizar en la actualidad que cuando Abagnale se encontraba al otro lado de la ley, según ha repetido en numerosas apariciones en medios.

El cibercrimen y los delitos informáticos proliferan a una velocidad tal que no hay margen para desarrollar formas de combatirlo. "Como ciudadanos y como usuarios esperamos que cada compañía provea servicios confiables y libres de ataque, pero esto probablemente no es factible", aseguran en el FTF. "Esto está llevando al **desarrollo de una gran industria en torno a la ciberseguridad**. En las compañías proveedoras se está convirtiendo en una especialidad y por otra parte emergen nuevos negocios basados en seguridad en la red", comentan. Incluso se puede llegar a convertir en un tema de Estado, como pudimos comprobar a finales de 2014 con el ataque informático a Sony Pictures Entertainment supuestamente perpetrado por el Gobierno de Corea del Norte como represalia ante el inminente estreno de la película *The Interview*, en la que se ridiculiza al líder de este país.

Sin embargo, la proliferación de empresas y especialistas en seguridad no evita que sigan



"El *big data* permite más acceso a todo, y su inteligencia se puede usar para el bien y para el mal", señalan en el FTF.

apareciendo nuevas y creativas formas delictivas en internet. Una cada vez más popular es el secuestro de datos, que consiste en cifrar archivos como fotos, videos o documentos valiosos. Ransomware, una variedad de malware de moda, se encarga de hacerlo automáticamente y de pedir el rescate de los datos por la vía digital. Quien quiera recuperarlos, deberá pagar. Y prevenirlo es complicado, ya que a menudo carecemos de herramientas para protegernos de estos ataques. **La vulnerabilidad aumenta en el caso de los dispositivos móviles.** "¿Es seguro tener datos en tu smartphone?", se preguntan en el FTF. La realidad es que no, y que la única manera 100% segura de evitar un ataque es no almacenar en él nada que no quiera que otros vean, o roben.

El fraude *online* no es el único que se multiplica gracias a la tecnología. Criminales, terroristas y narcotraficantes también se valen de las nuevas herramientas. "Desde el lado criminal usan los datos para profesionalizarse, unificar sus canales de comunicación y crear redes de criminales cada vez más grandes. La buena noticia es que hay nuevas y mejores herramientas para controlarlos, la mala es que ellos también pueden usar esas tecnologías. El *big data* permite más acceso a todo, y **su inteligencia se puede usar para el bien y para el mal**", señalan en el FTF.

La lucha entre policía y criminales es además desigual, como resaltan varios expertos del FTF. Dos variables otorgan ventaja a los criminales: la económica y la legal. "Las redes criminales manejan a menudo cantidades desorbitadas de dinero. Un traficante puede tener 20 billones de dólares en efectivo en su casa mientras el presupuesto total de la policía es de cuatro billones". A esto se añade

la ausencia de limitaciones legales, con las que sí cuenta la policía. "Usar *big data* sin límites facilita el desarrollo de las superorganizaciones criminales inteligentes y que siga creciendo su red y la cadena de suministro. Así que desde el lado de la policía el sistema legal tiene que cambiar para permitir, por ejemplo, el intercambio de datos e información entre todas las fuerzas de seguridad. De lo contrario, seguirán en una situación de injusta desventaja", señalan en el FTF.

Los expertos participantes en el debate también resaltan la necesidad de contar con una especie de **brigada *big data* contra el crimen**. Es decir, una unidad de análisis de datos que integre a científicos de datos en el cuerpo de policía. "En Nueva York de 49.000 policías solo hay 880 entrenados en datos. ¡Necesitan muchos más!", clama uno de los presentes. "Cada vez es más urgente una respuesta automática y ágil contra ciberdelincuentes y criminales y, si no cuentan con personal capacitado para ello, no podrán hacerlo", concluye.



Including a Photo of a Woman

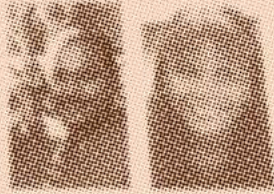
Including a Photo of a Woman increased response rate by as much as a 25% reduction in bounce rate (on a 1-month basis)



Photo of a woman
Photo of a man

Including a Photo of a Woman

Including a Photo of a Woman increased response rate by as much as a 25% reduction in bounce rate (on a 1-month basis)



mailchimp.com



Retos pendientes

4

Cap. 4/7

4.1 Qué medir y cómo librarse del ruido

4.2 Entender y aplicar

4.3 Propiedad versus privacidad versus seguridad



Internet **industrial:** ¿qué falta por hacer?

Stefan Bungart. Líder de GE Software Europe en GE Global Research

➔ **Mientras los fabricantes** siguen persiguiendo mayor productividad y rentabilidad, mejorando incesantemente la eficiencia y eficacia de sus operaciones, hay un tema concreto que ha llamado mucho la atención. Internet industrial, la conexión entre tecnología de producción y máquinas (tecnología operacional, OT por sus siglas en inglés) y los entornos informáticos dentro de la empresa y entre las empresas y sus colaboradores en la cadena de distribución ha sido el foco de atención en los últimos años.

Conectar así máquinas y equipamiento de producción promete distintos y valiosos beneficios. Acceso a más datos (y con más miga) sobre todos los aspectos de la operación y cada proceso individual, por ejemplo, o la capacidad de actualizar el firmware y el software de las máquinas con mayor frecuencia y periodicidad.

Los datos se pueden usar para entender mejor los parámetros de máquinas individuales como el desgaste habitual, y la combinación de datos de distintas máquinas sirve para entender mejor la interacción entre procesos y máquinas. No hay que ubicar todo el equipamiento en un mismo sitio para

lograr este beneficio, sino que se puede ubicar en plantas de fabricación distintas.

Las actualizaciones frecuentes de firmware y software permiten actualizar también, e incluso mejorar o sustituir las funcionalidades de las máquinas y el equipamiento. Las máquinas que se suelen cambiar así se denominan ordenadores definidos por software.

Internet en general y los protocolos de IP en particular permiten combinar máquinas y aplicaciones de la nube para alcanzar la productividad y rentabilidad que persiguen las empresas.

Hoy día, adoptar aplicaciones de internet industrial en fabricación se suele limitar a pilotos y pruebas de concepto. Los grandes fabricantes están explorando posibilidades. Hay fabricantes más pequeños cuya estrategia consiste en esperar y ver qué pasa. Cualquier ruptura en el proceso de fabricación o producción conlleva inmediatamente pérdidas de producción, y tiene un impacto en la cadena de distribución en general. En ciertas industrias, la pérdida de producción debida a estas rupturas alcanza cifras de millones de dólares por hora.

Las tecnologías de internet industrial, tanto hardware como software, se encuentran a día de hoy en una fase de desarrollo inicial, y por naturaleza son más vulnerables ante fallos y ataques. Para que acelere la adopción de internet industrial, es necesario hacer varias mejoras importantes.

Analítica y estrategias de datos

Los beneficios de las aplicaciones de datos individuales en mantenimiento y reducción de costes y el mayor tiempo en funcionamiento del equipamiento han quedado demostrados. En cambio, queda mucho por descubrir y entender sobre el uso de grandes cantidades de datos detallados en pro de la productividad y rentabilidad en la fabricación y cadenas de distribución. Esto requiere primero que los expertos en estrategia de datos sopesen bien los casos de uso y las aplicaciones, y que las tecnologías de análisis de datos extraigan y formen el conocimiento a partir de los datos.

Seguridad

Una máquina conectada a internet es vulnerable a ataques externos. Las estrategias de seguridad informática actuales son menos capaces de proteger los activos de producción críticos. Hace falta desarrollar nuevas estrategias y tecnologías de seguridad dirigidas concretamente a entornos de fabricación con tecnología operacional (OT).

Alta disponibilidad

Cuando se cae internet en una oficina, nadie se alegra, pero en la mayoría de los casos se puede seguir trabajando. Las redes no se construyen para tener disponibilidad 100% por temas de coste. Cuando se cae una aplicación de internet industrial, el coste que origina la pérdida de producción puede ser considerable. Los requisitos de producción son muy distintos y requieren distintas estrategias tecnológicas, tanto en OT como en tecnología informática que utilizan las empresas.

Apertura

El beneficio real de internet industrial es combinar máquinas y datos de distintos proveedores, en distintas ubicaciones, y posiblemente en distintas jurisdicciones. Muchos proveedores de equipamiento de producción utilizan formatos de datos y protocolos de comunicación y lenguajes registrados, y las máquinas no "hablan" entre sí. Las distintas jurisdicciones tienen distintos requisitos de protección de datos y de distribución transfronteriza de datos. Las aplicaciones de internet industrial deben ser flexibles para adaptarse a este entorno y a una tecnología en constante cambio. Los estándares abiertos y los estándares de industrias concretas pueden ayudar a mejorar la apertura.

Las empresas de fabricación siguen invirtiendo en aplicaciones de internet industrial y aprenderán a beneficiarse de estas aplicaciones. Los proveedores seguirán desarrollando hardware y software adaptado concretamente a las necesidades de internet industrial. Sin duda, habrá nuevos retos y habrá que resolverlos. Internet industrial no supondrá una revolución, sino más bien una evolución. Nos ofrecerá una nueva forma de plantearnos la fabricación y producción.

Introducción

4.0

➔ **Resulta que somos** cocineros y cae en nuestras manos un nuevo ingrediente que se llama *big data*. Todo el mundo habla maravillas de él y se dice que revolucionará la gastronomía. Lo primero que tenemos que hacer es ponerlo en contexto: enterarnos de dónde viene, a qué familia alimenticia pertenece, a qué sabe y con qué sabores puede tener un buen maridaje.

Como nos ha llegado en conjuntos de grandes volúmenes, el siguiente paso será limpiarlo para eliminar los datos basura y quedarnos con aquellos que realmente aportan sabor. Con esta selección debemos preparar un buen plato. Para ello contaremos con los mejores cocineros y deberemos convencer a los comensales de que les gustará probar nuestra lasaña de datos. Su presentación deberá ser atractiva, con las capas bien diferenciadas a simple vista.

Tras un rotundo éxito en la cocina, ahora que ya hemos disfrutado de nuestro elaborado manjar de datos se nos plantean varios dilemas: ¿los compartimos con otros? ¿Revelamos nuestra fórmula secreta receta para que otros la apliquen e incluso

mejoren a pesar de que eso signifique perder la exclusividad de nuestra receta?

A la vista está que **tenemos más grandes preguntas que grandes respuestas**, como bien aseguran los 'cocineros' del FTF. El drama, como subraya uno de los presentes, sería no ser capaces de darnos cuenta de los beneficios de nuestro gran ingrediente: el *big data*. Afortunadamente es una cuestión de actitud, por lo que depende de nosotros ganar o no la batalla ante los siguientes tres grandes retos identificados: una buena medición y limpieza de datos, una correcta aplicación y un equilibrio entre lo que compartimos y lo que no y cómo lo hacemos para no dañar la privacidad de los usuarios.

Qué medir y cómo librarse del ruido

4.1

¿Cómo llega una red social como Facebook a conocer nuestros sentimientos sin preguntarnos? Una vez más, la respuesta es el *big data*. A finales de 2014, un [estudio](#) publicado en la revista científica *Computers in Human Behavior* demostraba la posibilidad de conocer el estado anímico de los usuarios de esta red social con un 83% de precisión. Para ello, investigadores españoles de la Universidad Autónoma de Madrid han desarrollado una aplicación llamada SentBuk que hibrida de técnicas de análisis semántico y de aprendizaje computacional automático.

Gran parte del éxito de SentBuk reside en que **tiene en cuenta el contexto**: clasifica al usuario en función de variables como la polaridad de sus mensajes (previamente analizados), de los cambios emocionales detectados o de las reacciones a las emociones de sus amigos. Este tipo de método es exactamente el que los expertos reunidos en el FTF proponen como imprescindible para obtener datos valiosos. "El elemento humano es muy importante, no sirven los datos per se porque estos no son neutrales. Encontrar la manera idónea de cómo el algoritmo de estas herramientas debe interactuar con los seres humanos es un gran reto", sostiene **Juan Carlos López**, del Centro de Innovación de Roche.

"Entender el contexto y las relaciones causa-efecto es fundamental en este sentido", apunta **Emilio Méndez**, director del Centro de Nanomateriales Funcionales del Departamento de Energía de Estados Unidos. Por tanto, lo relevante del



Juan Carlos López ▲
Líder de Relaciones Académicas y
colaboración de Roche Innovation Center



Emilio Méndez ▲
Director del Center for Functional
Nanomaterials del U.S.



trabajo con grandes volúmenes de datos empieza antes incluso de procesarlos. Por una parte, hay un **margen de error que tiene que ver con la intencionalidad real** del usuario al compartir sus mensajes. "El volumen y el valor operacional y transaccional versus los datos declarativos. O, dicho de otra manera, lo que la gente hace frente a lo que dice, postea o tuitea", subrayan en el FTF.

La idea de monitorización pasiva sin contexto es incompleta. No podemos perder el componente emocional del comportamiento.

Por otra parte, varios participantes en el FTF señalan que importa más la calidad de los datos por sí mismos que cómo se analizan. "Virtualmente todos los datos son muy descriptivos, Necesitamos que permitan inferencias causales para entender por qué esos datos de superficie son los que son. Sin contexto pueden ser alarmantes o dar falsas impresiones".

De falsas impresiones habla **Antonio Damasio**: "Los dispositivos desarrollados hasta el momento son muy simples y su nivel de error es considerable. Por ejemplo, la fiabilidad de las mediciones de dispositivos para controlar la presión sanguínea dependen del contexto, el dato en sí no es nada si no sabemos si cuando se produce me estoy levantando o estoy en reposo. Las mediciones dependen de cómo te sientes en términos de bienestar, salud, etc. en cada momento", explica el neurólogo.

Esto es exactamente a lo que se refería **Vicki Seyfert-Margolis** en su aportación al grupo de trabajo sobre médicos y pacientes: "**La idea de monitorización pasiva sin contexto es incompleta. No podemos perder el componente emocional del comportamiento** porque refleja lo que la gente realmente necesita y quiere usar y cómo conecta con otras personas".

Así que falta contexto y sobran datos... Y, al igual que los cocineros, los científicos de datos también deben tirar a la basura los datos sucios o insípidos. Esto representa la parte más importante y más laboriosa de su trabajo. Es lo que Méndez define como "separar la señal del ruido, lo importante de lo no importante, lo que es relevante de lo que no". Es decir, distinguir entre datos finos y datos espesos o profundos. La tarea es ardua porque no hay limitaciones en términos de cuántos datos podemos recopilar. Se trabaja con volúmenes enormes y esto puede ser un arma de doble filo. "Tenemos que pensar acerca de cuántos datos son demasiados porque la realidad es que no tenemos ni idea", indica López.

Una vez tenemos los datos necesarios, la siguiente pregunta -como apunta Chris Meyer- es: "¿De esos datos, cuánto necesitamos saber? La respuesta



Carla Brodley

Decana de la College of Computer and Information Science de la Northeastern University

Los expertos del FTF coinciden en la necesidad de contar con más personal formado en la ciencia de datos y aseguran que es capital para el aprovechamiento del poder del *big data*.

determinará cómo se realizará la limpieza de datos. "Una vez más nos enfrentamos a un reto: cómo nos aseguramos de que realmente estamos limpiando bien los datos y no estamos trabajando con ruido. De ello depende la veracidad de nuestros resultados", retoma López. En su opinión, la calidad de los datos "se mide por cuántas predicciones podemos hacer con ellos". **El poder de la bola de cristal.** "El valor del *big data* reside en crear subconjuntos de datos sobre lo que nos están indicando los conjuntos iniciales que hemos obtenido", sostiene el directivo de Roche.

Carla Brodley añade un comentario al respecto: "Los informáticos odian el término *big data* porque está vacío de contenido". Y es que los datos por sí mismos no prueban su valor. "Hay compañías que perciben una necesidad de datos y hay una tensión entre lo que se necesita y lo que dichas organizaciones están haciendo al respecto", añade Maartin den Braber. "Estamos trabajando en redes profundas que serán factibles si somos capaces de incluir estos procesos dentro de la directiva de grupo que establece las diferentes configuraciones de los usuarios que acceden al sistema informático de la organización", completa Brodley.

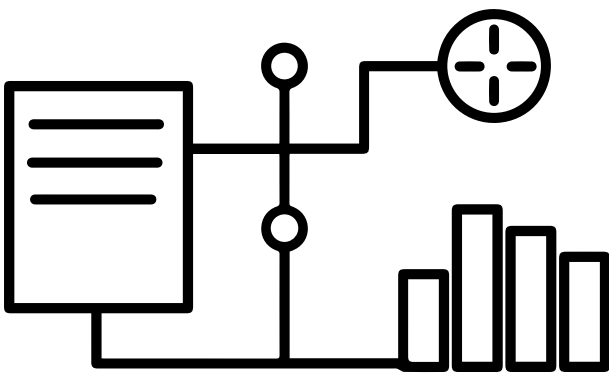
"La falta de estándares y formatos unificados para los datos es un obstáculo para este fin. **Los datos son desordenados, incompletos, inconsistentes, y se presentan en diferentes formatos**, por lo que su gestión se convierte en una habilidad clave. Requiere de una estrategia tanto a nivel de corporación como de negocio. Y también de una infraestructura de tecnologías de la información adecuada que afronte de manera integral el desarrollo de plataformas de predicción, su integración como KPI a nivel financiero, la familiarización y formación del personal y los aspectos jurídicos sobre los derechos de los datos", sintetiza **Stefan Bungart**.

Otra cuestión relativa a la apuesta de las compañías y que es clave para la correcta selección y limpieza de datos es la carencia de recursos humanos capaces de realizar esas tareas de forma óptima. Los expertos del FTF coinciden en la necesidad de contar con más personal formado en la ciencia de datos y aseguran que es capital para el aprovechamiento del poder del *big data*.

Entender y aplicar

4.2

La utilidad de las herramientas de visualización para transmitir información densa, difícil de entender, reside en que la información visual aporta claridad y orden.



Los profesionales de datos no solo son necesarios para extraer la información útil de los datos y deshacerse de la que no sirve. También son fundamentales para entenderlos y transmitirlos a quienes deban aplicarlos ya sea con fines de mejora de rendimiento interno, de desarrollo de negocio o de una transformación más amplia. "Lo que realmente importa aparte de que hayas recopilado los datos correctos es que puedas explicar los resultados a la persona o personas dentro de la organización que van a usarlos", indica Brodley. De lo contrario, podemos tener la información delante y no saber leerla.

Tal y como señalan en el FTF, este es el drama de la carencia de **entendimiento del valor de los datos**. "Tienes que encontrar ese valor y si tú no tienes esa

habilidad tienes que encontrar a gente que lo haga. No se trata solo de una cuestión técnica, la aplicación del análisis de datos masivos para impulsar los resultados requiere una nueva experiencia comercial que a menudo no está disponible dentro de la organización. En ese caso será necesario contratarla externamente y potenciar las habilidades del personal interno con un amplio programa de aprendizaje", detalla **Stefan Bungart**. "Solo así conseguiremos historias de éxito que aumenten la inversión, y estas comienzan donde el valor es más fácilmente identificado y capturado. No es ninguna broma, esto puede derivar en cientos de millones de dólares en ahorros o en ingresos derivados de un solo cliente", añade el directivo de General Electric.

En síntesis, se trata de transformar el negocio a la era de los grandes volúmenes de datos mediante el aprovechamiento de sus beneficios en términos de eficiencia y de reducción de costes. Y pasa por externalizar la complicada labor profesional a consultores externos para ayudar a acelerar el proceso de cambio. Los participantes en el FTF sostienen que este tipo de medidas son necesarias para detectar dónde deben poner atención los empleados.



Miguel Arias
COO en CartoDB

Las **herramientas de visualización** complementan esta tarea. Las imágenes son mucho más fáciles de procesar y muy útiles frente la saturación de información que a menudo produce el trabajo con grandes volúmenes de datos y que puede hacer sentir abrumados a los no especialistas. La utilidad de las herramientas de visualización para transmitir información densa, difícil de entender, reside en que la información visual aporta claridad y orden.

Por eso es interesante su aplicación, entre otros, en informes de resultados de análisis *big data*, lo cual facilita la comprensión de los resultados estratégicos. **"Ayuda a que los profesionales de los negocios entiendan el *big data*",** asegura **Miguel Arias**, director de operaciones de CartoDB. Lo dice con conocimiento de causa, ya que su empresa se dedica a crear visualizaciones de grandes volúmenes de datos con proyectos en áreas tan distintas como la salud, la educación, banca y finanzas, inteligencia de negocios, gobernanza, arquitectura o periodismo. Y para muestra un botón:

Imaginemos que somos los responsables de *marketing* del Real Madrid y queremos conocer nuestro impacto en Twitter durante la final de la Champions League. Imaginemos que en lugar de un largo texto explicativo en el que se describe el tráfico de tuits geolocalizados en cada país y región, tenemos un mapa interactivo en el que lo vemos todo de un vistazo, y con la posibilidad de elegir el nivel de detalle que queramos solo haciendo zoom. Esto es exactamente lo que hizo CartoDB durante la final de la Champions League la primavera de 2014 en Lisboa entre el Real Madrid y el Atlético de Madrid. Y [este es el resultado](#): un claro ejemplo de uso

eficaz de las herramientas de mapeado y visualización para una comprensión efectiva de los datos.

Sin embargo estas herramientas no son suficiente... **Taryn Sullivan** pone el dedo en la llaga: "Tenemos preciosas analíticas y sofisticados gráficos pero necesitamos saber qué hacer con ellos y cómo aplicarlos". Para Sullivan este es el gran reto, especialmente en países en vías de industrialización como China. "Hay que adentrarse en los datos y actuar", remarca. Y hay que hacerlo sin prisas, a juicio de Lita Sands. Su experiencia dirigiendo la transformación digital de Novartis la ha enseñado que **estas transformaciones requieren años**. "Lo hicimos muy bien al principio, creando un clima para el cambio, pero nos equivocamos al pensar que lo podíamos hacer en tres meses".

Sands recuerda que, si bien en ese tiempo lograron implementar los cambios técnicamente (dotando a la fuerza de ventas de tabletas y pasado el papel a formatos digitales) no fue así humanamente. "Es un viaje de años, no es tan fácil como preguntarte si quieres que pase y que al momento ya esté hecho. Es importante que las organizaciones entiendan que tienen que construir estas capacidades en la organización y revisar cada año su plan y sus mensajes para sostenerlo y reforzarlo continuamente. Y la manera en que apreciamos el rendimiento a nivel corporativo a menudo es de cuatrimestre en cuatrimestre", explica.

Hay diferentes enfoques a la hora de abordar este tipo de procesos. El aprendizaje multitarea es según Brodley "la mayor tendencia ahora, que acapara un tercio de los artículos científicos publicados sobre

*"El único **big data** en el que la gente está interesada son los datos personales".*

Hacer los datos interpretables, integrar a los humanos en el aprendizaje y realizar dicha integración de manera responsable, preservando la privacidad.

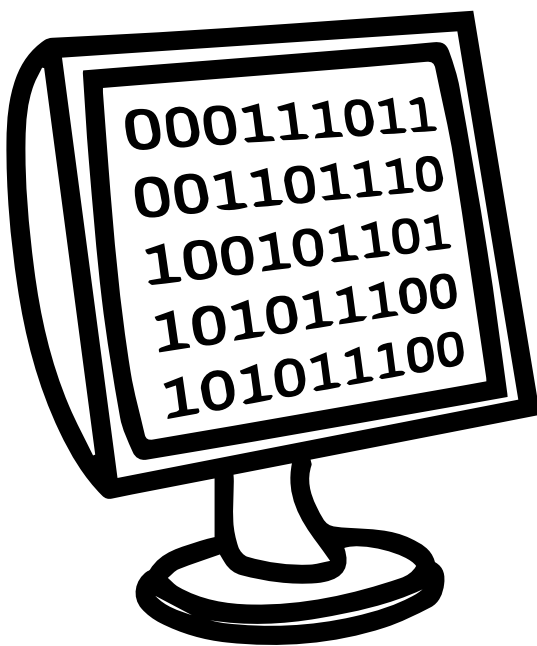
procesos de aprendizaje computacional automático". "Pongamos por caso que analizo un modelo de un hospital (A) que funciona y lo intento aplicar a otro hospital (B). No lo puedo hacer automática e inmediatamente por las diferentes prácticas, los tratamientos, etc. entre ellos pero tampoco quiero tener que recopilar todos los datos del nuevo hospital (B) porque es un proceso caro. Así que lo que hago es tomar el modelo que he aprendido del hospital A, coger unos pocos más datos del hospital B y transferir ese conocimiento", ejemplifica Brodley para explicar el modelo de aprendizaje multitarea.

Otro debate sobre la aplicación de los resultados del análisis de datos más allá de la industria es su utilidad a nivel humano. Según Den Braber, "el único

big data en el que la gente está interesada son los datos personales". Se refiere, por ejemplo, a obtener patrones de su comportamiento, "no necesariamente para actuar en consecuencia sino simplemente para conocerlos". Por tanto, un gran reto en este sentido es cómo las creencias de las personas son modificadas por las nuevas tecnologías y el acceso a la información. Es decir, cómo extraer el potencial beneficio en aspectos como la salud pública que aportan los sensores y aplicaciones para la medición de actividad que muchas personas ya usan.

Los expertos del FTF también se preguntan cómo se podrían integrar los datos de esas aplicaciones y de otras más específicas para el seguimiento de ciertas patologías en el día a día del paciente con los datos tomados en los hospitales. "Si soy un doctor que se enfrenta a todos estos datos, ¿qué hago con ellos? ¿para qué los uso? ¿para mejorar el abordaje del paciente, para aprender sobre lo que funciona y lo que no del sistema? Si estoy recopilando datos de pacientes, ¿dónde paro?", se pregunta Brodley.

La respuesta no está clara, aunque sí el objetivo: "Tenemos que utilizar estos datos para reinterpretar nuestras medidas clínicas actuales para un éxito en términos de mejora de la calidad percibida de la vida humana, de la salud y de los resultados que importan a los pacientes en relación con su capacidad para funcionar", señalan en el FTF. En opinión de Brodley, para ello es importante "**hacer los datos interpretables, integrar a los humanos en el aprendizaje y realizar dicha integración de manera responsable, preservando la privacidad**". Este último es precisamente otro de los grandes retos que enfrenta el *big data*, como veremos a continuación.



Propiedad *versus* privacidad *versus* seguridad

4.3

"Te vigilan. El Gobierno cuenta con un sistema secreto que te espía cada día, cada instante. Lo sé porque yo mismo lo creé. Lo hice para detectar actos terroristas pero la máquina lo ve todo: violentos crímenes que implican a personas ordinarias como tú. (...) Víctima o delincuente, si tu nombre aparece, te encontraremos". Quien lo asegura no es uno de los expertos reunidos en el FTF sino el protagonista de la serie [Person of interest](#).

Esta situación no es parte de la realidad (todavía) pero muestra el inmenso poder de los sistemas de videovigilancia y de reconocimiento facial combinados con la minería de datos. Este tipo de tecnologías -aunque no tan avanzadas como la de la serie- ya existen. De hecho, Reino Unido anunció el pasado verano de 2014 que la policía del condado de Leicestershire iba a empezar a usar un nuevo software de reconocimiento facial llamado NeoFace. Este software compara las mediciones tomadas a partir de la imagen de una cara con las 92.000 que contiene la base de datos de la policía de Leicestershire, y lo hace según ellos "con una elevada tasa de éxito", [tal y como recoge la BBC](#).

La mayoría de las imágenes con las que NeoFace compara sus mediciones proceden de las cámaras de videovigilancia instaladas ya en numerosos rincones de las ciudades. Su presencia no deja de ser polémica y ha generado acalorados debates sobre el **derecho a la intimidad y a la anonimidad de los ciudadanos**. Por una parte, está la privacidad de las personas y la seguridad de sus datos, que se



Retención de los individuos y las empresas a compartir datos.

contrapone a su seguridad física cuando interfiere en una acción policial que podría beneficiarse de disponer de dichos datos.

Aquí entra en juego también el concepto de propiedad y su delimitación. ¿A quién pertenecen los datos que Gobiernos, fuerzas de seguridad y otras organizaciones recogen de las personas?

¿Dónde empieza el derecho de unos a usarlos y acaba el derecho de otros a la protección de su propia información? "El acceso a los datos es una gran preocupación, al mismo tiempo que lo es la privacidad. Por tanto aquí hay un conflicto de intereses, y el abuso de los datos es una gran preocupación. ¿A quién debemos gritar "no seas malvado"? ¿A los gobiernos, a las corporaciones, a las fuerzas militares, a los medios de comunicación, a todos ellos?", se pregunta en voz alta uno de los participantes en el FTF.

Las implicaciones del *big data* relativas a la propiedad, la privacidad y la seguridad son por tanto innegables. "El **manejo responsable de datos** es la parte que creo que requiere de más trabajo de la comunidad de informáticos para preservar la privacidad de los consumidores, especialmente la relacionada con los pacientes. La mayoría de la gente no entiende nada cuando hacen clic y dicen "sí, me gusta este comentario de Facebook" y no se dan

Identificar al 90% de más de un millón de consumidores anónimos tan sólo analizado algunos datos de las operaciones realizadas con sus tarjetas de crédito.

cuenta de qué es lo que esto implica exactamente en relación con la privacidad. Incluso aunque lo entiendan no se preocupan ni tienen precaución porque su afán por usar la tecnología supera las intromisiones en su privacidad que pueda acarrear. Pero no creo que este sea el caso con los datos médicos. Necesitamos crear algoritmos que permitan el manejo de datos al tiempo que se preserve la privacidad", reflexiona Brodley.

El ámbito de la salud no es el único que merece especial atención. Los datos financieros, en especial los provenientes de tarjetas de pago, son especialmente sensibles. Así lo ha demostrado un equipo de investigadores del Instituto Tecnológico de Massachusetts (MIT) que ha logrado **identificar al 90% de más de un millón de consumidores anónimos tan sólo analizado algunos datos de las operaciones realizadas con sus tarjetas de crédito**. La revista científica *Science* se ha hecho eco de este estudio en un especial titulado "[El fin de la privacidad](#)".

A la luz de esta información, no es extraño que existan compañías como Abine que cuenta entre sus servicios con un sistema denominado [DeleteMe](#) que se dedica a borrar información personal de la red. "Está claro que este es ahora el reto del *big data*: la privacidad, la individualidad y la anonimidad y cómo abordarlas teniendo en cuenta que su valor procede del uso de diferentes fuentes y de la integración de los usuarios en el proceso", señalan en el FTF.

Esteban Moro cuestiona quién es propietario del valor de los datos que emanan de los ciudadanos. En su opinión, "este valor es tan grande que la sociedad se tiene que beneficiar de ello". "Naciones Unidas ya cuenta iniciativas que hacen a las compañías compartir datos para casos de emergencia", comenta. Se refiere al programa [Global Pulse](#) que promueve el conocimiento de las oportunidades del *big data* para la ayuda al desarrollo mediante alianzas de intercambio de datos de los sectores público y privado.

Entre otras cosas, Global Pulse utiliza datos en tiempo real, seguimiento de redes sociales, teléfonos móviles, blogs y plataformas de comercio en línea



Andreu Veà
Presidente de la Internet
Society (ISOC-ES)

para obtener indicadores que pueden alertar de malestar social o de un posible desastre natural de forma anticipada. El objetivo es preparar un plan preventivo adecuado que, según Naciones Unidas "podría salvar vidas".

Los participantes en el FTF creen que, a pesar de programas como este, la **reticencia de los individuos y las empresas a compartir datos** -ya sea por cuestiones de privacidad o por motivos competitivos o económicos- en un entorno de cambio regulatorio y de intercambio de valor actualmente incierto **impedirán la materialización de muchos de los beneficios y promesas del big data en nuestra sociedad.**

Al respecto, uno de los presentes en el FTF señala que "hay buenas intenciones pero mala regulación". "La gente en las organizaciones no entiende cómo proteger los datos a la vez que permite la innovación. Hay una carencia de entendimiento del valor de compartir datos entre verticales de la industria, lo cual tiene enormes beneficios para todo el mundo", añade. Un paso más allá van quienes como Moro y como **Andreu Veà** -presidente de la Internet Society y Digital Champion de la Comisión Europea para España- consideran necesario garantizar un sistema de datos abiertos para el bien público, más allá del mero intercambio empresarial. Arias y varios de los presentes están de acuerdo, y hablan de datos masivos de calidad accesibles a todo el mundo. "Si no, corremos el riesgo de que a pesar de estar ahí los datos no sean usados por nadie", puntualiza Arias.

El problema es doble: de acceso y de regulación. De acceso porque este sigue siendo muy fragmentado

y desigual: **no todo el mundo tiene las mismas posibilidades de obtener los datos**, ya sea por cuestiones económicas, tecnológicas o legales. Y de regulación porque no hay un marco que establezca las reglas del juego. "¡Todo el mundo tiene miedo a la regulación!", exclama uno de los presentes en el FTF. De ahí que muchos coincidan en que las empresas deben empezar a discutir con los afectados el derecho de propiedad de los datos "antes que sea el Gobierno quien lo haga".

La inseguridad legal al respecto va en aumento, según Bungart, que está convencido de que la complejidad jurídica en este sentido también crecerá. Y, por si el problema de la recolección de datos no parece lo suficientemente complejo, hay quien apunta que "es, de lejos, mucho mayor de lo que podemos apreciar". El reto del equilibrio entre propiedad, seguridad y privacidad corre, por tanto, el riesgo de convertirse en un obstáculo para el aprovechamiento del *big data*. Una barrera legislativa, como veremos en el próximo capítulo.

Impedirán la materialización de muchos de los beneficios y promesas del *big data* en nuestra sociedad.



La llave al futuro

5

Cap. 5/7

- 5.1 Legislación y marco legal
 - 5.2 Silos
 - 5.3 Formatos, estándares y calidad
 - 5.4 Factor humano y resistencia
 - 5.5 Habilidades y recursos
 - 5.6 Mercados infradesarrollados
-



La llave al futuro

J.C. Herz. Diseñadora y consultora estratégica con Big Data

➔ ¿Hasta qué punto soy dueño de mi comportamiento y todos sus resultados? ¿Cuándo deja de ser mía una transacción social o comercial?

Un debate en profundidad sobre tecnología a menudo saca a relucir tensión y soluciones intermedias que poco tienen que ver con los sistemas técnicos, y mucho con nuestros valores individuales y colectivos. El entusiasmo o el miedo de la gente ante la tecnología y lo que ésta pueda traer, refleja una profunda convicción en lo que respecta al papel del individuo en la sociedad: si el individuo es el actor principal cuya libertad y acciones deben preservarse por encima de todo, o si el bien común o comunitario requiere implícitamente una participación pública. Los debates sobre la privacidad, las ciber-amenazas, la visibilidad que gobiernos y empresas comerciales puedan alcanzar del uso que hace la gente del sistema técnico, y la promesa de las analíticas de *big data* en el mundo de la salud pivotan en torno a siguiente pregunta: "¿Hasta qué punto soy dueño de mi comportamiento y todos sus resultados? ¿Cuándo deja de ser mía una transacción social o comercial?" Un individualista pedirá autoridad para tomar decisiones y participar en el valor económico que genera su "descarte de datos digitales". Un corporativista o colectivista considera que los datos anonimizados de particulares son un recurso natural que se puede excavar y refinar.

Estas preguntas normativas son las que dan empuje a decisiones jurídicas y legislativas, a prácticas

empresariales y a estándares industriales. Si la información es poder, la decisión de compartirla—o el requisito de compartirla—es básicamente una decisión política, sea una cuestión política, literalmente, o un tira y afloja entre empresas o incluso dentro de una misma empresa. La decisión de "compartirás los datos" entre hospitales y clínicas refleja la voluntad política de un sistema sanitario nacional (o en caso de Estados Unidos, de un gobierno pagador), mucho más que los aspectos técnicos de dicha cuestión. Las peleas y eternizaciones en estándares industriales o interfaces de programación de aplicaciones (API) giran casi siempre en torno a conflictos sobre recursos, ventajas competitivas, puntos de apoyo a favor de avances disruptivos frente a la posición atrincherada de quienes vienen disfrutando de la posición dominante. El hecho de que este teatro kabuki en torno a los datos se extienda a través de distintos ámbitos (medios de comunicación y comercio electrónico, finanzas y sanidad, transporte, educación, seguridad nacional y el proceso de buen gobierno en sí mismo) recalca que las visiones encontradas son el resultado de sistemas de valores encontrados. Conforme el paisaje tecnológico va cambiando, nos da un ejemplo de cómo van cambiando paradigmas muy profundos, que no son tecnológicos y que son mucho más difíciles de acelerar que el hardware, el software o las API.

Introducción

5.0

➔ **Stephen Hawking** ha sido "el primer científico serio que se ha planteado la posibilidad práctica de los viajes en el tiempo", según asegura en *El universo en una cáscara de nuez*. Si alguien lo lograra, "podría cambiar la historia y dominar el mundo -dice el científico- pero "solo unos pocos de nosotros somos lo suficientemente locos para trabajar en un tema tan políticamente incorrecto".



Jessica Bland
Investigadora de Tecnologías
del Futuro de Nesta

Alocados o no, los participantes en el FTF se adentran en un viaje al futuro del *big data* del que no podrá escapar quien se atreva a seguir leyendo. Conocer la respuesta a grandes preguntas sobre esta ciencia está ahora a su alcance. Porque, junto a los retos de los que ya hemos hablado, se presentan una serie de obstáculos legales, humanos, de mercado, de formatos, de acceso y de habilidades. Y la pregunta

es: ¿Seremos capaces de superarlos? ¿Arribaremos a la tierra prometida donde se hace realidad el potencial del big data? ¿Llegaremos todos, o solo unos pocos?

Teniendo en cuenta la situación actual, se plantean varios posibles hitos para un futuro inmediato y a medio y largo plazo:

Legislación y marco legal

5.1

La nueva regulación aprobada se centraría de manera estricta en combatir el fraude y el robo de identidad.

A Esteban Moro le ilusiona pensar que en un futuro cercano -tanto como 2016- gobiernos, empresas y agencias de las naciones unidas se verán empujadas a llegar un acuerdo de **compartición de datos para luchar contra grandes problemas como la pobreza, el acceso a los alimentos, las epidemias o el crimen organizado.**

Desde la perspectiva de las compañías, la diseñadora de producto **J.C. Herz** cree que ese mismo año Apple marcará de nuevo el camino, en este caso el de la privacidad, la seguridad y la confianza del consumidor en las aplicaciones móviles. ¿Cómo? Con un cambio de dinámica "más poderoso que cualquier legislación" que consiste en requerir a cada *app* presente en su tienda iTunes que revele cómo y qué datos recoge de los usuarios, y cómo los monetiza.

Estos datos se mostrarían en la clásica pantalla "acerca de" en la página de la iTunes Store. "Una especie de Charity Navigator [una organización sin ánimo de lucro que evalúa las organizaciones benéficas en Estados Unidos] donde vemos a quién le estamos dando el dinero" asegura Herz. En su opinión, el verdadero valor de esto reside en que "si se convierte en una norma social a través de la cual todo el mundo entiende fácilmente lo que se está obteniendo de tu *app* móvil, **los datos toman parte en la ecuación del consumidor**".

Con una visión más proteccionista lo mira **Ana María Arboleda**. La directora de Atlantis Healthcare en España apunta que, para estas fechas, todo tipo de negocios (ya sean grandes, medianas o pequeñas empresas) contarán con una Política de

La desconfianza generada por la invasión y entrada a todo tipo de datos llevará de regreso a una mentalidad conservadora caracterizada por la creación de células aisladas de datos y conectividad.



J.C. Herz
Diseñadora y consultora
estratégica con Big Data



Ana María Arboleda
Directora General de Atlantis
Healthcare Spain

Protección de Datos, algo que **Jessica Bland** cree que se generalizará en 2017 con una **regulación europea de protección de datos**. Según esta investigadora de nuevas tecnologías de Nesta, dicha normativa "armonizará el acceso a los datos y la legislación de la privacidad a nivel regional, y posiblemente sea aceptada como un estándar global". **Maarten Den Braber** cree que este estándar de protección de datos podría estar integrado en las leyes de todos los países en 2023, "siendo optimistas".

Volviendo unos años atrás, hasta 2018, veremos la llegada de una gran ayuda al control de problemas de privacidad de la mano de la informática, con algoritmos de aprendizaje automático que prevenirán los problemas de privacidad, según **Carla Brodley**. A pesar de ello **Steve Trachtenberg**, presidente emérito de la Universidad George Washington y patrono de la Fundación Innovación Bankinter, vaticina que los políticos -en concreto en el Congreso de Estados Unidos- serán cada vez más temerosos acerca de la dirección que está tomando

el *big data* y verán "implicaciones orwellianas en la regulación del uso de la información".

Esto será en 2020, cuando los nativos digitales ya estén en el poder y desarrollen todo tipo de medidas contra las barreras del *big data*, tal y como prevé **Taryn Sullivan**. A pesar de ello, un acto masivo de pirateo terrorista tirará abajo la web en 2020. La desconfianza generada por la invasión y entrada a todo tipo de datos llevará de regreso a una mentalidad conservadora caracterizada por la creación de células aisladas de datos y conectividad, según vaticina uno de los presentes.

Miguel Arias también cree que algo parecido puede suceder, aunque sitúa la acción cinco años después, en el año 2025: "Una campaña global orquestada por supercriminales destruirá internet y tendremos que remodelar completamente la legislación y la forma en que pensamos sobre el *big data*". Arias cree que, en tal caso, **la nueva regulación aprobada se centraría de manera estricta en combatir el fraude y el robo de identidad**.

Silos

5.2

Según el diccionario de la Real Academia Española un silo es un "lugar subterráneo, profundo y oscuro". En las catacumbas del *big data* se encuentran la igualdad de acceso a los datos y a las herramientas que permiten procesarlos, estructurarlos, analizarlos y darles uso. **Julia Li** cree que el primer hito en este sentido será realidad en 2017. Tiene que ver con la predisposición de las ciudades y de las empresas, que trabajarán en la **construcción de un ecosistema de datos**, al tiempo que los individuos toman una actitud más abierta hacia esta ciencia y sus implicaciones.

Otra tarea para la industria en este sentido - que **Chris Meyer** calcula que se acometerá en 2019- será la de establecer unos estándares y una taxonomía para las APIS de programación. En la misma línea pero un año después sitúa Arias la democratización del acceso a herramientas de datos que ahora poseen las grandes compañías. Cree que esto hará "que los científicos de datos y la gente hablen el mismo idioma".

Probablemente como consecuencia de ello, Moro predice que en 2025 **"todo el mundo estará haciendo big data con sus teléfonos móviles** con acceso a datos abiertos y a herramientas de creación colaborativa". "Para nosotros hoy en día trabajar con volúmenes masivos de datos es un problema porque necesitamos una infraestructura sobre la que manejarnos pero para entonces ya contaremos con ella", asegura. Es más, se atreve a dar un titular: "Un grupo de niños encuentra un potencial tratamiento para una enfermedad olvidada usando herramientas libres de manejo de datos y de visualización abierta y colaborativa durante un *hackathon*".

No parece ir muy desencaminado Moro, teniendo en cuenta que una niña de 7 años aficionada a la

tecnología ya ha sido capaz de piratear una red wifi pública en 11 minutos. Ese fue el resultado de un [experimento](#) conducido por la empresa de seguridad Hide My Ass! del que medios de comunicación de todo el mundo se han hecho eco.

La construcción de un ecosistema de datos, al tiempo que los individuos toman una actitud más abierta hacia esta ciencia y sus implicaciones.

Formatos, estándares y calidad

5.3

En estrecha relación con el acceso a una infraestructura de trabajo con datos masivos que según los expertos del FTF saldrá de la sombra del silo como tarde antes de 2021 está el acceso a los datos en sí mismo. Según **Julie Freeman**, artista y directora del programa de Cultura del Arte en el Instituto de Datos Abiertos (ODI), esto ocurrirá antes de liberalizarse las herramientas. Será en 2017 cuando "los certificados de datos abiertos empezarán a adoptarse a nivel mundial y los formatos se irán alineando", afirma Freeman.

En 2019, dos años después, sitúa **Juan Carlos López** la universalización de un **formato común para los historiales de los pacientes** "que pueda ser compartido por todos los científicos y facilitar un cuidado interactivo". "Esto pasará cuando las empresas se den cuenta de lo que supone en términos de reducción de costes y las personas sean conscientes de los beneficios que puede aportarles, de manera que unos estarán motivados a hacerlo en base a crear eficiencias y los otros por sus propias necesidades", explica el directivo del Centro de Innovación de Roche.

Esto llevará, según López, a la firma de un acuerdo internacional para acordar un estándar sobre las normas que deben observarse para informar sobre la salud y los datos de los pacientes. Será en 2021 y, un año después (2022), "tendremos un **marco legal sobre qué datos de los pacientes deben ser compartidos por las empresas**".

Un poco antes se habrá alcanzado otro gran hito: en 2021 **toda la tierra usará un único estándar de datos**, según **Andreu Veà**. "Pasará como hoy en día



Julie Freeman ▲
Artista y líder de Datos del Culture Art Programme en el Open Data Institute (ODI)



Richard Kivel ▲
Emprendedor e inversor en Tecnología médica y biofarmia y Patrono de Fundación Innovación Bankinter

con el TCP [Protocolo de Control de Transmisión], que se ha convertido en un gran facilitador". De este nuevo estándar pasaremos a una nueva tecnología que "permitirá la comprensión y análisis de datos de alta calidad frente a datos de baja calidad". Esto ocurrirá en el 2024 según **Richard Kivel**, inversor y patrono de la Fundación Innovación Bankinter.

Ese mismo año podremos disfrutar ya de una de las tecnologías más esperadas: los **coches conectados**. Muchos equipos de investigación de todo el mundo y grandes compañías como Google llevan años trabajando en ello, y Kivel cree que el aumento de datos de alta calidad y de la disponibilidad de sensores avanzados permitirá el uso de estos coches que se conducirán solos. "Esencialmente veremos una transformación en la que serán los proveedores de servicios -y no las personas- los que posean la mayoría de los vehículos en circulación, y no habrá necesidad de seguros, *parkings* y ese tipo de cosas. Usaremos los coches como cualquier otro producto alquilable que te lleva, te deja, se va y vuelve a por ti o manda a otro a recogerte", prevé el inversor.

Factor humano y resistencia

5.4



Edén Shochat

Fundador de Aleph y Patrono de Fundación Innovación Bankinter

¿Qué pintan las personas en todo esto? ¿Cómo influye su predisposición a favor o en contra de contribuir al avance de esta ciencia? Los hechos demuestran que hay motivos para desconfiar a la hora de proporcionar datos a terceros. A pesar de ello se da la paradoja de que esta desconfianza resulta aligerarse cuando al usuario le interesa mucho acceder a una tecnología concreta o por el efecto de arrastre de la masa. Es el caso, como hemos visto, de redes sociales como Facebook.

Sin embargo, esto no evita que la desconfianza hacia estas aplicaciones siga creciendo. Por eso **Edén Shochat** -fundador de la empresa Aleph que ayuda a emprendedores israelíes a construir negocios escalables y patrono de la Fundación Innovación Bankinter- ve a Google y a Facebook colaborando en 2016 "en la creación de una *start-up* que desarrollará un **nuevo sistema federado de identidad** para el análisis de datos basado en permiso y que aumentará el gráfico de Facebook".

Otro acuerdo con Google de por medio es el que predice **Joseph Kvedar**. Esta vez la otra parte

será Apple. En 2019 ambos gigantes tecnológicos acordarán la creación de una sencilla e intuitiva interfaz para que los consumidores puedan compartir datos individuales de sus almacenes privados de datos. Ese mismo año otro factor facilitará aún más la compartición: **la conectividad será global**, con acceso para todo el mundo.

Esto "incrementará las capacidades de proporcionar datos, que estarán disponibles en cantidades cada vez más voluminosos para que el mundo pueda trabajar mano a mano", según Freeman. Otro capacitador más para los usuarios de a pie será la disponibilidad de herramientas de aprendizaje computacional para uso personal. Brodley afirma que esto sucederá en 2020 y "cerrará la brecha entre los científicos y la gente corriente" gracias a "una interfaz fácil de usar y accesible para cualquiera".

También se centra en el autoaprendizaje el vaticinio de otro experto presente en el FTF, que habla de la aparición de "**nuevas disciplinas educativas** en torno a la ciencia de datos y su relación con el comportamiento humano". En su opinión, estas surgirán "facilitadas por el aprendizaje remoto y autocertificado y al margen del sistema de universidades, colegios, etc.", lo cual "revolucionará la forma en la que funciona la educación hoy en día".

Pero toda moneda tiene dos caras y el acceso a herramientas sencillas que capaciten a todo el mundo para analizar datos llevará años más tarde -en 2025- a que muchos analistas de datos se queden sin empleo, según Li. "Estas herramientas se volverán un juguete aburrido y volverá a ganar valor la intuición humana, minimizándose el valor del *big data*".

Habilidades y recursos

5.5

Emergerá un campo líder en las escuelas de negocios y universidades.

La situación de desempleo generalizado de analistas de datos que describe Li sería, no obstante, una consecuencia de un proceso a largo plazo y que tiene como fecha de comienzo este mismo año: 2015. "Es ahora cuando se están definiendo las competencias y requerimientos para ser un buen analista de datos", asegura esta experta en aprendizaje experimental. En su opinión, no se trata ya de científicos sino de una figura más cercana a los negocios, el comercio electrónico, la gamificación y la ciencia del comportamiento.

"Estas habilidades en conjunto ayudarán a entender mejor los datos, y **la necesidad de contar con personas preparadas para ello de manera inmediata provocará que emerja un campo líder en las escuelas de negocios y universidades**, que contarán con toda una línea de programas de preparación intensiva y exprés", asegura Li. Como resultado, en dos años (2017) alguna de las universidades líderes a nivel global será conocida por contribuir a formar expertos en análisis de datos enfocados en *e-commerce* y gamificación.

Otra manera de verlo -aunque compatible- es la que tiene Sullivan. La fundadora de EEx cree que para 2018 **la ciencia de datos "estará integrada en cada disciplina de estudio** y esto permitirá a más individuos

imaginar cómo crear valor con la información disponible y dar vida a un diseño movido por la interacción humana con los datos". Con una visión parecida Brodley predice que para 2019 "todos los estudiantes universitarios habrán estado expuestos en algún momento a la ciencia de datos".

Tal vez debido a esta interacción **Wilfried Vanhonacker** augura un desengaño. "Nos daremos cuenta de que **el big data no es la maravilla que todos creíamos**". Habrá buenos usos pero no al nivel de lo que esperamos en este momento de sobreexcitación que vivimos a fecha de 2015", señala Vanhonacker. "Como científico de datos suelo decir que un buen experto en esta ciencia puede crear cualquier dato y conducirlo a cualquier cosa, así que mejor seamos conscientes de ello", afirma con sorna.

Puede que la perspectiva de Vanhonacker cambie de hacerse realidad la idea de **Emilio Méndez**, que confía en que para 2019 "contaremos con una aplicación desarrollada por IBM que permitirá **la extracción del 1% de información útil en cualquier conjunto de datos** y desechar automáticamente de los datos inútiles".

Si esto aumenta el potencial de explotación del *big data* tendría sentido que se sigan poniendo esfuerzos en la formación de personas que trabajen con datos. Entonces podría suceder que, como afirma *Vicki Seyfert-Margolis*, Estados Unidos, China y Europa firmen en 2022 un acuerdo conjunto para una campaña de educación en *big data*: una política abierta "que permita el intercambio de talento e incremente la presencia de recursos humanos valiosos en todo el mundo". Y tal vez esas personas usen **ordenadores cuánticos**, que -según Shochat- en 2025 estarán ya disponibles para ser aprovechados por la ciencia de datos.

Mercados infradesarrollados

5.6

Hablábamos en el capítulo *Show me the Money* de la constatación de que hay un mercado para los datos. Pero a la hora de vislumbrar el futuro en este sentido son pocos los participantes en el FTF que aportan ideas. **Puneet Batra**, científico de datos y cofundador de la *start-up* LevelTrigger, es uno de ellos. Su visión a medio plazo (para 2018) pasa por el desarrollo de **un marketplace donde se podrán comprar y vender datos entre empresas**.

Los dos factores clave para permitirlo son, a juicio de Batra, "un entendimiento por parte de quien compra datos de qué significa eso y cómo va a usarlos" y "un conocimiento por parte de quien los vende de las implicaciones que tiene en términos de privacidad". "Eso conducirá a un mercado de datos más fluido", indica.

Algo más disruptivo plantea **Iyad Rahwan**, profesor asociado en el Instituto Masdar y en proceso de traslado al laboratorio Media Lab del MIT. Rahwan es un experto en encontrar agujas en el pajar de los datos y lanza convencido un titular para 2025: **"Un algoritmo totalmente automatizado gestiona con éxito una ciudad durante un mes entero"**. Sus labores serán operacionales: pagará a la gente, encontrará problemas, distribuirá recursos... Todo ello a partir de *big data*, sensores y de las herramientas disponibles. Lo único que, en opinión de Rahwan, no hará el algoritmo, será tomar decisiones estratégicas. Sin embargo, deja caer que "la gente se empezará a cuestionar qué tipo de gobierno podría existir si esto se escala a otras formas de toma de decisiones".

Estos dos son los hitos para el futuro de la comercialización de datos que ven los expertos del FTF. Meyer opina que tal vez la carencia de aportaciones refleja una falta de confianza en el valor del intercambio de datos. Sin embargo, cree que el verdadero motivo es que las **oportunidades se van a crear a partir de los cambios estructurales** anteriormente descritos, "particularmente sobre regulación, acceso y formato de los datos". "Podemos esperar otro tipo de reacciones necesarias en estos ámbitos, como que las compañías empiecen a hacer las cosas fáciles y que el *big data* sea algo que cualquiera pueda entender".

Son solo algunas de las posibilidades que los expertos del FTF contemplan para un futuro en el que individuos, corporaciones, gobiernos e instituciones deberán **trabajar juntos sobre la base de la transparencia y la confianza para superar las barreras del big data**. Y también, por qué no, para hacer del mundo un lugar mejor. Si los datos, como dice Freeman, demuestran que las predicciones sobre asuntos como el cambio climático son reales, "todo el mundo tendrá que centrarse en cambiar la forma en la que hace uso del planeta y del medio ambiente".

Claro está, de igual modo, que las cosas pueden ir a peor gracias al uso de la ciencia de datos y de las sofisticadas herramientas basadas en ella si se utilizan con fines delictivos, terroristas y criminales. Ya se sabe: la tecnología no es mala ni buena, todo depende del uso que se haga de ella. El futuro no está en sus manos, pero sí en las nuestras.

El futuro del Big Data

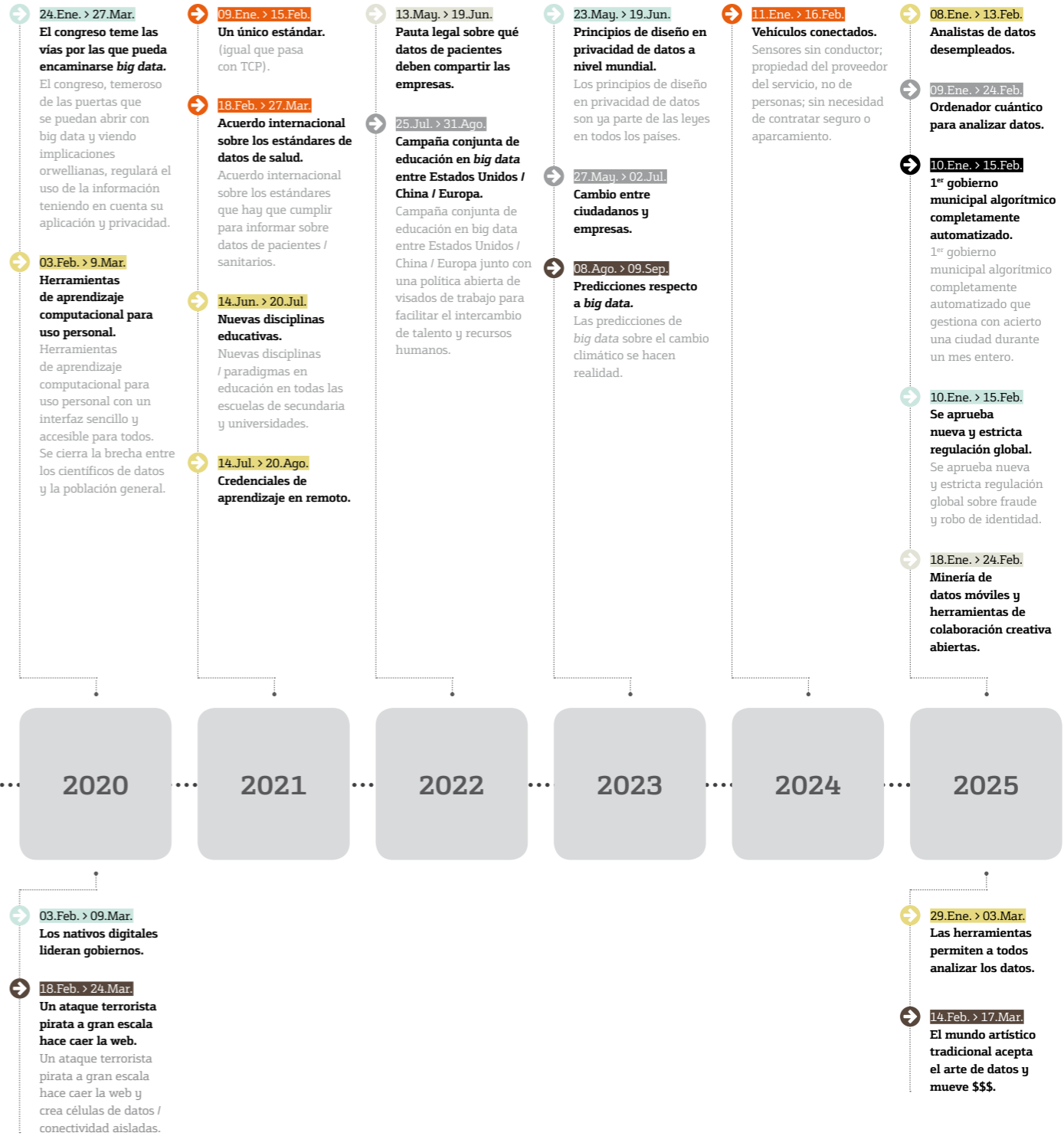


El futuro del Big Data

Time line hitos destacados



- ➔ Hitos en mercados por desarrollar
- ➔ Hitos en formatos, calidad y estándares de datos
- ➔ Hitos en relación con el factor humano y la resistencia que se pueda oponer
- ➔ Hitos en legislación y marco legal



- ➔ Hitos en aptitudes y recursos
- ➔ Hitos respecto a los compartimentos estanco (Acceso, API)
- ➔ Otros hitos

Glosario

6

- > **1. Analíticas en tiempo real o *real time big data analytics*:** es la capacidad de las nuevas herramientas tecnológicas de *big data* para analizar grandes conjuntos de datos en el mismo momento en el que se están generando dichos datos.
- > **2. Analíticas predictivas o *predictive big data analytics*:** se generan como resultado de la extracción y análisis del conocimiento histórico que se encuentra en los datos masivos y señalan posibles patrones, resultados o tendencias futuras.
- > **3. API o *Application Programming Interface*:** la Interfaz de Programación de Aplicaciones es el punto de contacto a través del cual un informático puede acceder a una aplicación sin necesidad de conocer su funcionamiento, usando el lenguaje informático.
- > **4. Aprendizaje computacional automático o *automatic machine learning*:** modelo de autodidáctica computacional basado en inteligencia artificial y que se realiza a partir de análisis comparativos de datos estructurados, semiestructurados y no estructurados. A menudo requieren interacción humana por parte del analista y se nutren de ella.
- > **5. Aprendizaje multitarea:** es un modelo aprendizaje computacional automático que analiza de forma conjunta varios problemas o tareas relacionadas para extraer los puntos comunes y los más útiles de todas ellas y así mejorar la situación de partida.
- > **6. *Back-office*:** en términos informáticos el *back-office* es el software que procesa información de la empresa como bases de datos, registros, operaciones, inventarios, etc. y que se encarga del mantenimiento de esta información de trastienda.
- > **7. *Big Data*:** es el conjunto de tecnologías, técnicas y herramientas que hacen posible la recogida, procesamiento y análisis de volúmenes masivos de datos, y también la visualización de los resultados. El propósito es convertir la información hallada en esos grandes conjuntos de datos en algo útil como estadísticas, patrones de comportamiento, análisis de rendimiento, etc.
- > **8. *Crowdsourcing*:** técnica que consiste en implicar a las masas en la búsqueda soluciones a problemas complejos. En el ámbito corporativo requiere abrir los datos necesarios para el trabajo de las personas ajenas a la organización. El *crowdsourcing* es un recurso de aplicación de la innovación abierta.
- > **9. Cuadro de mando o *dashboard*:** pantalla que presenta todas las métricas clave del negocio -los KPI más importantes- en un solo lugar. Esboza una imagen del panorama general con respecto a esos indicadores y facilita la detección de tendencias.
- > **10. Datos basura o *dirty data / thin data*:** son grandes volúmenes de datos inútiles que no aportan información de valor en un contexto de análisis *big data*.

- > **11. Datos limpios** o *clean data / thick data / deep data*: son aquellos datos de calidad que determinarán los resultados del análisis *big data*.
- > **12. Directiva de grupo** o *Group Policy Object* (GPO): sistema de configuración de la infraestructura informática que sigue cada organización y que determina lo que los usuarios pueden hacer en su equipo dentro del entorno computacional.
- > **13. Gamificación**: se trata de la aplicación de técnicas de la teoría de juegos a la interacción con diferentes públicos (internos y externos) de la organización y que sirve para atraer y para motivar a dichas personas. Algunos ejemplos de aplicación son la incorporación de recompensas para impulsar la acción ante un reto propuesto, indicadores de estado y de actividad con los amigos o compañeros, tablas de clasificación y posibilidad de subir de nivel, datos de rendimiento, barras de progreso, etc.
- > **14. Hackathon**: se trata de una especie de maratón intensiva de programación que se lleva a cabo durante un elevado número de horas seguidas, con el objetivo de aportar soluciones -en forma de prototipo de API- a uno o varios retos propuestos.
- > **15. Holocracia**: es un sistema de organización que distribuye la autoridad en grupos autónomos o círculos que identifican, gestionan y resuelven los objetivos y retos de forma independiente aunque de forma integrada con el resto de la organización.
- > **16. Indicadores clave de rendimiento** o *Key Performance Indicators* (KPI): son parámetros medibles que fijan los objetivos de la organización y que sirven para medir y controlar el desempeño a nivel estratégico y operativo de una empresa al comparar su evolución con dichos parámetros. De esta manera se identifican deficiencias y se visualiza el potencial de mejora de procesos de negocio.
- > **17. Innovación abierta** u *open innovation*: es un tipo de innovación que se basa en un proceso de integración de personas y comunidades ajenas a la organización en su cadena de innovación. Por tanto, aprovecha los recursos externos para encontrar de forma más rápida y ágil nuevas y mejores soluciones a los problemas y retos corporativos.
- > **18. Inteligencia de negocios** o *Business Intelligence* (BI): es el conjunto de métodos para mejorar la toma de decisiones de negocio mediante el uso de sistemas de apoyo basados en hechos. Aplicado al *big data* se puede definir como la técnica a seguir para la toma de decisiones basadas en datos aplicada a los negocios.
- > **19. Intraemprendedor**: término acuñado recientemente para definir a aquellas personas emprendedoras dentro de la organización, que no solo son proactivas y aportan ideas sino que las llevan a cabo: desde la concepción hasta su materialización.
- > **20. Malware**: es un tipo de software malicioso diseñado para causar daños informáticos.
- > **21. Manejo responsable de datos** o *responsible data mining*: es una forma de usar los datos que respeta la privacidad de las personas de las que se han obtenido.
- > **22. Minería de datos** o *data mining* es la ciencia que se dedica a encontrar las agujas en el pajar del *big data*, que extrae lo valioso de entre la ingente cantidad de información para obtener lo valioso.

- > **23. Redes profundas** o *deep networks*: sistema de desarrollo computacional que trata de imitar el funcionamiento de las redes neuronales profundas para lograr que una máquina pueda detectar, reconocer, recordar y responder como la mente humana. Su manera de aprender se basa en representaciones de datos y en modelos que aprenden de estas representaciones.

- > **24. TCP** o *Transmission Control Protocol*: el Protocolo de Transmisión de Control es un sistema que utiliza protocolos comunes para que diferentes sistemas informáticos puedan intercambiar información a través de internet.

- > **25. Toma decisiones basadas en datos** o *data-driven decision making*: es la aplicación práctica del conocimiento obtenido a partir del informe de resultados del análisis *big data*. Es el último paso del proceso, después de haber realizado la extracción de valor y el análisis de los datos. Si este muestra, por ejemplo, cambios en los patrones de compra de los consumidores y prevé ciertas tendencias, la toma de decisiones consistirá en actuar para adaptarse a dichos cambios y anticiparse a la materialización de las tendencias previstas.

Recursos

7

7.1 Informes y estudios

- > Informe *Are you thinking too small about big data?* de IBM sobre recoge las previsiones de crecimiento del valor del *big data* para 2015: <http://www.ibm.com/smarterplanet/us/en/smarter-enterprise/perspectives/big-data-and-analytics.html>
- > Informe de resultados financieros de 2014 de la *start-up* Splunk: http://files.shareholder.com/downloads/AMDA-RWDLH/3981313229x0x657273/5C7E9352-D951-4F14-BE2E-6B513C307C05/Splunk_Annual_Report_FY14.pdf
- > Informe de resultados financieros de 2013 de la empresa Acxiom: <http://d3u9yejw7h244g.cloudfront.net/wp-content/uploads/2013/09/2013-Annual-Report.pdf>
- > Estudio publicado en la revista científica *Computers in Human Behavior* sobre la posibilidad de conocer el estado anímico de los usuarios de Facebook con un 83% de precisión a través de análisis *big data* de los mensajes e interacciones de los usuarios: www.sciencedirect.com/science/article/pii/S0747563213001751
- > Compendio de estudios y artículos publicados en el especial *The end of privacy* de la revista científica *Science* sobre las oportunidades y los dilemas que plantea el uso de los datos que los usuarios arrojan a internet y sobre la vulneración de la privacidad y la anonimidad de las personas: www.sciencemag.org/content/347/6221/490.full?intcmp=collection-privacy

7.2 Artículos y noticias

- > Artículo donde se habla del avance que podría permitir en un futuro que todos los datos del mundo puedan ser almacenados en un disco duro de ADN del tamaño de una cuchara: <http://singularityhub.com/2015/02/20/worlds-data-could-fit-on-a-teaspoon-sized-dna-hard-drive-and-survive-thousands-of-years>
- > Artículo sobre un análisis realizado por la empresa Jawbone a partir de los datos sobre la calidad del sueño tomados de las personas expuestas a un mismo terremoto: <https://jawbone.com/blog/napa-earthquake-effect-on-sleep>
- > Artículo sobre la capacidad de predicción del *big data* acerca de nuestro perfil psicológico: http://elpais.com/elpais/2015/01/12/ciencia/1421084469_835718.html
- > Artículo sobre la capacidad de predicción del *big data* acerca de las personas que nos importan: www.technologyreview.com/news/533536/att-builds-an-assistant-app-with-social-skills
- > Noticia sobre el uso policial de un software de reconocimiento facial de criminales llamado NeoFace que ya se emplea en Reino Unido: www.bbc.com/news/uk-england-leicestershire-28307938
- > Artículo sobre un experimento conducido por la empresa de seguridad Hide My Ass! que tuvo como resultado el pirateo de una red wifi por parte de una niña de 7 años: www.dailymail.co.uk/sciencetech/article-2919762/Hacking-Wi-Fi-child-s-play-Seven-year-old-shows-easy-break-public-network-11-minutes.html

7.3 Sitios web

- > Sitio web de la comunidad Quantified Self, creada para conocer las últimas novedades en herramientas y aplicaciones de automedición y para ayudar a la gente a entenderlas: <http://quantifiedself.com>
- > Sitio web de la empresa SHIFT, cofundada por Maarten den Braber, experto presente en el FTF: <http://shiftbsp.com>
- > Sitio web de la pulsera inteligente Fitbit para la medición de actividad diaria: <https://www.moves-app.com>
- > Sitio web de la pulsera inteligente UP de Jawbone para la medición de actividad diaria: <https://jawbone.com/up>
- > Sitio web informativo sobre el libro *The Victory Lab* del analista político Sasha Issenberg, en el que explica cómo el Partido Demócrata usó el big data para ganar las elecciones presidenciales de 2008 y de 2012: www.thevictorylab.com
- > Sitio web de la competición GE Flight Quest de General Electric: www.gequest.com
- > Sitio web de la empresa My Own Med fundada por Vicki Seyfert-Margolis, experta presente en el FTF. Se trata de una plataforma en la que paciente, cuidador y profesionales sanitarios comparten una misma interfaz, tienen acceso a los mismos datos y pueden trabajar en la misma página usando una arquitectura en nube.
- > Página web de servicio DeleteMe que se dedica a borrar información personal de la red y que forma parte de la cartera de productos de la empresa Abine: www.abine.com/deleteme/landing.php

- > Sitio web del programa Global Pulse de Naciones Unidas para promover el conocimiento de las oportunidades del *big data* en la ayuda al desarrollo mediante alianzas de intercambio de datos de los sectores público y privado: www.unglobalpulse.org

7.4 Aplicaciones y plataformas

- > Plataforma web Kaggle, una comunidad de científicos de datos enfocada en la resolución colectiva de problemas a través de competiciones abiertas: www.kaggle.com
 - > Portal de gobierno abierto de Canadá: <http://open.canada.ca/en>
 - > Portal de gobierno abierto de Estados Unidos: www.whitehouse.gov/open
 - > Portal de gobierno abierto del País Vasco: <http://www.irekia.euskadi.eus/es>
 - > Portal de gobierno abierto de la provincia de Zaragoza: www.zaragoza.es/ciudadania/gobierno-abierto
 - > Portal de datos abiertos de España: www.datos.gob.es
 - > Aplicación Moves para la medición de actividad diaria: www.moves-app.com
- ### 7.5 Otros
- > Gráfico de la frecuencia cardíaca de un hombre mientras hace una propuesta matrimonial: <http://imgur.com/mbOPX2L>
 - > Vídeo de apertura de la serie *Person of interest*: <http://youtu.be/WOnQ8CD3v4g>



