

BIG DATA

El poder de convertir
datos en decisiones

A un clic de las TIC

Telefonica

Big data: el poder de
convertir datos en
decisiones

Índice de contenido

Introducción	4
Vivir en un mar de datos	5
<i>Big data</i> , una marea imparable	6
<i>Big data</i> e IoT: un mar de oportunidades para las grandes empresas	8
Datos al servicio de la persona	10
Sectores pioneros en <i>big data</i>	13
Banca	
Cómo facilita <i>big data</i> la transformación hacia una banca digital	14
Salud	
<i>Big data</i> sanitario: el acelerador del conocimiento y la decisión clínica	17
<i>Big data</i> en Sanidad, clave en la creación de tratamientos eficaces y eficientes	20
Smart cities	
<i>Big data</i> : piedra angular de las <i>smart cities</i>	24
Sector agrícola	
<i>Big data</i> y la nueva revolución del sector agrícola	27
Educación	
<i>Big data</i> en las aulas: del “café para todos” a una “educación a la carta”	30
Deporte	
<i>Big data</i> en el deporte: ¿dopaje tecnológico o mejor toma de decisiones?	33
Periodismo	
El periodismo de datos como creador de historias	36
Datos: de las estadísticas y las barras a historias para la sociedad	38
Autores	40
Julián García Barbosa	41
Roberto García Esteban	41
Félix Hdez. de Rojas	42
Visitación López López	42
Mercedes Núñez	43

Introducción

Se dice que revolucionará la salud como **acelerador del conocimiento y la decisión clínica y ayudará a la creación de** tratamientos realmente eficaces y eficientes. Es la piedra angular de las *smart cities*, clave en la transformación hacia la banca digital y en una nueva agricultura de precisión. Ha dado vida al periodismo de datos, ha cambiado la forma de ver el deporte, se ha convertido en una necesidad crítica para los negocios por la ventaja competitiva que representa...

En este momento todo el mundo está hablando de *big data*. No es que sea el futuro, sino que ya está en nuestras vidas. Quienes lo ven como el término de moda aún no han entendido su magnitud: “**Data is the new oil**” pero, como el petróleo, es necesario encontrar, extraer, refinar, distribuir y monetizar esos datos. Se trata de convertirlos en información, ésta en conocimiento y finalmente en decisiones.

El valor de *big data*, que trabaja con datos anónimos y agregados, **reside en encontrar patrones que ayuden a anticipar el futuro: prever crisis económicas, prevenir enfermedades**, trazar la ruta idónea del transporte público o acertar al abrir un local en una determinada zona.

Big data puede aportar valor en todos los sectores y ayudar a mejorar la gestión de todo tipo de servicios. *Big data*, por tanto, para *big problems*, con tres grandes ámbitos de aplicación:

- Mejorar la solución a problemas que ya tienen respuesta
- Solucionar problemas que no tienen respuestas
- Encontrar problemas que no sabíamos que teníamos

Pero la recopilación y almacenamiento de datos exige una formulación previa del problema y, desde luego, un análisis posterior. Recopilación y almacenamiento se automatizan pero encontrar el algoritmo que extraiga la información de utilidad implica un recurso humano: el *data scientist* o experto que traduce los datos obtenidos en conocimiento y acción, y ahí es donde reside realmente el valor. Hace ya unos años la revista Harvard Business Review la calificó como “la profesión más sexy del siglo XXI”. De hecho, por lo complejo del perfil, hay un crecimiento exponencial de ofertas de trabajo en este sentido. Entre otras capacidades, requiere dotes de computación, conocimiento matemático/estadístico, así como del negocio y del problema que se aborda. Por tanto, desentrañar la verdad que los datos esconden requiere siempre una perspectiva, una mirada, un relato humano.

Por su impacto en la economía, el empleo y la innovación, *big data* está ya en la agenda de los Gobiernos; incluso hay un consorcio público privado, cPPP, para integrar a todos los actores de *big data* en Europa porque hace falta una estrategia conjunta para ser competitivos en el espectro mundial (se habla ya de una “brecha del *big data*”).

También plantea múltiples retos: de almacenamiento, de gestión, cuestiones éticas y legales, de financiación, negocio versus privacidad... ***Big data* genera un gran poder pero a la vez conlleva una gran responsabilidad.**

Apasionante como este eBook recopilatorio de *posts* sobre *big data* publicados en *A un clic de las TIC*. ¡Esperamos que disfrutéis de su lectura!

Vivir en un mar
de datos

Big data, una marea imparable

Félix Hdez. de Rojas (@Felixhderojas)



Vivir en un mar de datos puede representar no solo una ambiciosa aventura tecnológica, sino también un [jugoso beneficio](#) para nuestro país o, más bien, para toda Europa. Los entendidos cifran que *big data* proporcionará un crecimiento del 1,9 por ciento del PIB en 2020 con 4,4 millones de empleos ya directamente involucrados en 2015.

De ello se habló en la primera sesión del ciclo de conferencias “[Mar de datos](#)” que Fundación Telefónica celebró en el segundo trimestre de 2015 y que nos permitieron ampliar nuestros conocimientos sobre *big data* e IoT.

En la actualidad se refieren al **dato como el petróleo digital del s. XXI**. El llamado oro negro fue en su momento motor de riqueza y transformación: ciertamente nadie recuerda los campos despoblados de Texas, pues ahora la industria petrolífera ha aupado esta región como una de las diez economías más prósperas y avanzadas del planeta. **Pero** el petróleo por sí solo no basta, **necesita de** otras industrias anexas: la de la transformación y sus refinerías... E igual sucede con *big data*, como se puso de manifiesto en este encuentro. Esto implica, entre otras cosas, **un entorno regulatorio propicio** que, sin embargo, no violente las libertades personales, un mercado útil para dichos datos, **un ecosistema innovador** de pequeñas y medianas empresas que desarrollen algoritmos y sus aplicaciones, sin olvidar una **infraestructura de**



investigación bien organizada y con financiación.

De todo ello se habló y los ponentes expusieron sus puntos de vista con brillantez. Ciertamente las **aplicaciones en banca, telecomunicaciones, sanidad e investigación** son referentes y demuestran actuales **casos de éxito**. Pero, más allá del bombo que se le da a *big data*, **ya muchos se manifiestan prudentes** debido a los costes de los actuales dispositivos y sensores, pues no son lo bastante baratos y sus baterías aún no duran lo suficiente, entre otros motivos. En realidad, como dijo Bruno Cendón, de TST Sistemas, **nuestros dispositivos IoT son aún como los móviles de segunda generación de principios de 2000**, ¿y quién los recuerda?

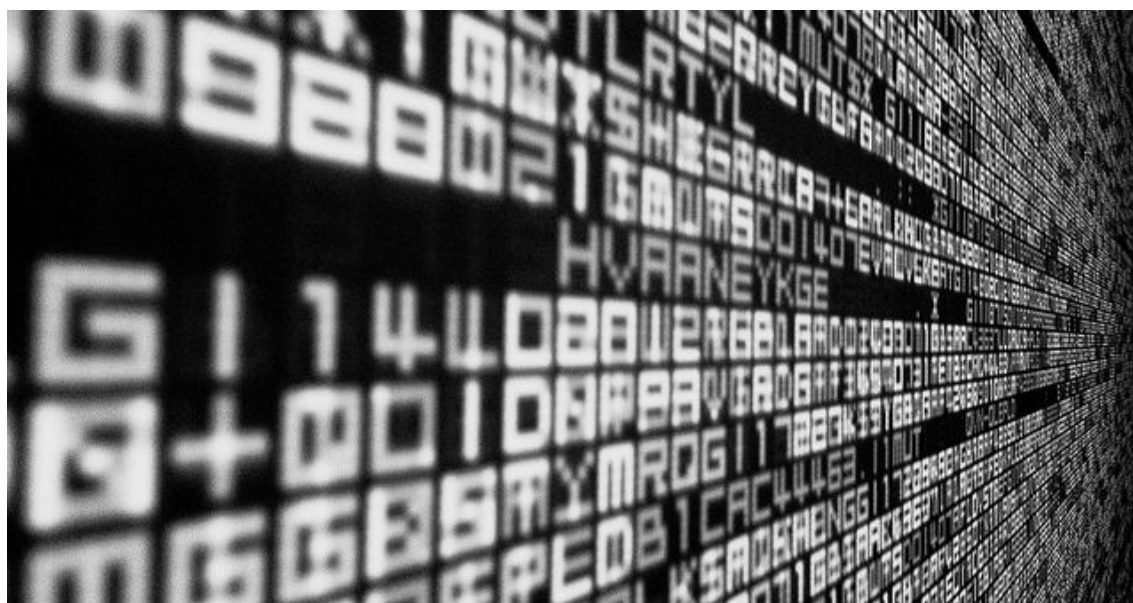
Pese a las dificultades existentes (quién no las encuentra cuando uno es pionero) quisiera terminar con un mensaje positivo: **los datos son la gran oportunidad para una gestión eficiente, son fuente de transparencia ciudadana y de avance del conocimiento**. Evidentemente no son la panacea (a este respecto hubo alguna interesante reflexión sobre la dualidad cuantitativa-cualitativa de la información), pues en cualquier análisis numérico **correlación no implica causalidad**... aunque la falta de datos significa siempre la más completa oscuridad y cerrazón al progreso.

Por eso, **¡encendamos la luz! Arriba el telón, enchufemos nuestros sensores a la nube y que los algoritmos matemáticos se abran paso en la sinrazón**. Lo peor de todo es no saber lo que sabes... y ya lo dijo Sócrates, aunque al revés, hace muchos siglos: solo sé que no sé nada. ¡Sócrates era un admirador del *big data* a su manera!



Big data e IoT: un mar de oportunidades para las grandes empresas

Félix Hdez. de Rojas (@Felixhderojas)



La segunda sesión de 2015 del [ciclo “Mar de datos”](#), en Fundación Telefónica, estuvo dedicada a las grandes empresas y a las [aplicaciones de mercado](#) de *big data* e Internet de las cosas. Esta vez los ponentes, siempre de lujo, fueron (y cito por orden de aparición) [Lucía Álvarez](#), [Francisco Manzanero](#), [Elena Pisonero](#), Mónica Martínez Walter y [David del Val](#).

Permitidme meter en una coctelera las ideas que se presentaron y, así, en precipitado desorden, resumirlas para despertar vuestro interés y que visitéis el [vídeo](#) del encuentro.

Lo primero, y como buen punto de partida, no hay que confundir un “*big data*” *problem* con un *big “data problem”*. Puede parecer de Perogrullo y, sin embargo, no lo es en muchas ocasiones. Lo cierto es que ahora vemos grandes oportunidades y posibles aplicaciones donde muchos solo tienen escasa eficiencia y un mal uso o desperdicio del tesoro de sus datos. Hay ya **grandes casos de éxito: valoración de riesgos en compañías de seguros**, curvas de demanda para agentes del sistema eléctrico, **análisis de patrones de consumo en aguas**, **personalización de ofertas en grandes superficies comerciales**, **análisis del lenguaje natural aplicado en medicina** y en la lucha contra el cáncer; pero también se mencionaron grandes oportunidades en el mercado de [prevención del fraude financiero](#), la [agricultura intensiva](#) y en muchos



ámbitos de nuestra administración central (¡los impuestos, Dios mío!) o también en la defensa del país...

Aunque me encantó saber que *Clash of clans* (este videojuego de estrategia para móviles es mi debilidad) había utilizado **tecnología big data para predecir el desenganche de los jugadores** (ahora entiendo mucho de lo que ha pasado en mi tableta en los últimos meses), ya había leído en el libro *Lean Analytics* de Alistair Croll y Ben Yoskovitz cómo **las métricas de negocio y el emprendimiento pueden aprovecharse del big data con éxito**.

Retos hay miles, ya se ve, pero, como decía Baltasar Gracián en su “Manual de la prudencia”, sé que es mejor dejar algunas preguntas sin responder. Así que aquí os arrojo unas cuantas sobre las que se debatió con gran interés en la sesión: **¿Cómo ayudar a las empresas a extraer el valor del dato? ¿Cómo democratizar el acceso a éste? ¿Cómo capturar todos los datos destacados y presentarlos al negocio en tiempo real y con patrones relevantes? ¿Cómo analizar con eficacia los datos no estructurados? ¿Cómo manejarnos con las innumerables fuentes de información y con el volumen de información que traerá el IoT? ¿Cómo tratar los nuevos requisitos de velocidad exigidos para las nuevas redes y la exigencia en la duración de las baterías de los cacharros? ¿Cómo se desarrollarán tecnologías habilitadoras para conectar los objetos al mar de datos? ¿Tendremos en el mercado masivo europeo estándares de ciudad** que hablen de IoT y de *big data*?

Evidentemente, para los anteriores desafíos necesitaremos de un nuevo profesional: un matemático-estadístico-programador, un perfil nuevo que llaman ya **big data analytics scientist** que, sin embargo, sin ser de letras, debe tener la suficiente creatividad como para entender y defender un negocio y, ante todo, debe ser capaz de comunicar estos hallazgos, dotarlos de solvencia. Es decir, un matemático con olfato, capaz de destilar los datos para sacarles jugo. Estos perfiles se los rifan en algunas universidades y ya están apareciendo cursos para formarlos con excelencia.

Para concluir, el público, en el turno de preguntas, nos premió con nuevos interrogantes: ¿Cómo congeniar un dato abierto y accesible con un uso ético del mismo? ¿Cómo sostener nuestro pequeño mundo de **privacidad en este mar de datos**?

Y se dijo entonces que “el individuo es el centro de los datos” (sic). Estoy de acuerdo y, como diría Galileo delante del tribunal inquisidor cuando espetó «*Eppur si muove*», la tecnología o el mercado habrán de apartarse y ceder su valor frente **al ciudadano, que es quien realmente importa**, se pongan como se pongan. **De cualquier otra forma, creo que este mar de datos tan interesante se convertiría en tormenta de libertad, vacía de progreso.**



Datos al servicio de la persona

Félix Hdez. de Rojas (@Felixhderojas)



La [tercera y última sesión de la edición de 2015 del ciclo “Mar de datos”](#), de Fundación Telefónica, una cita que es un “*must go*” dentro del panorama tecno-cultural madrileño, estuvo enfocada a poner en valor las **nuevas aplicaciones de *big data***. En esta ocasión los ponentes fueron Sebastián Lancestremère, director general de Microsoft; el doctor Julio Mayol, catedrático de la Universidad Complutense; Carlos del Cacho, *chief data scientist* en Job & Talent; Pedro Diezma, CEO de Zerintia; David Barrientos, responsable de comunicación de ANFAC (Asociación Española de Fabricantes de Automóviles, Turismos y Camiones) y mi compañero Antonio J. Sanchez, responsable de IoT de Telefónica.

Deportes, [wearables](#), medicina, recursos humanos, nuestros vehículos... Todo conectado. El menú fue inmejorable, con **ejemplos específicos del valor del dato** en estos mundos tan diferentes. Intentaré resumir los aspectos más reseñables.

Comenzamos con el **caso de éxito del Real Madrid**, que me fascinó sobremanera por su ambicioso objetivo, tan brillante y masivo: **acercar la experiencia del espectáculo deportivo más allá del propio estadio físico a 450 millones de seguidores, la mayoría de ellos fuera de España**. Para ello Microsoft y Sebastián Lancestremère proponían usar dispositivos tan originales e innovadores como las [gafas Microsoft HoloLens](#). Sobre el tema de los *wearables* hay otros muchos ejemplos prácticos de tecnología



“para llevar puesta” y de forma “invisible” (es decir, que busca ante todo pasar inadvertida a nuestros ojos y se integra con naturalidad en nuestro cuerpo), según palabras de Pedro Diezma. Pero, además, **pronto llegará una tecnología** más allá de la ropa, las pulseras o los relojes, una tecnología que superará el problema de las baterías (¿quizá porque los nuevos dispositivos se cargarán solos?) y **que, en cierta medida, nos hará “superhombres”**. Me encantó en este sentido el ejemplo que Pedro puso de la ropa que nos recuerda los **abrazos de nuestros seres queridos...**

Y respecto a la salud, Julio Mayol expuso e insistió en el propio **valor de los datos médicos** y nos inundó de preguntas: ¿para qué (nos) sirven los datos? ¿Cuánto cuesta explotarlos? ¿Y almacenarlos? ¿Qué problemas de seguridad conllevan? Explicó, así mismo, la **necesidad de un modelo sanitario que incremente los servicios pero que ante todo proporcione ventajas al paciente...**

Al margen de la salud, también se habló de personas y mucho... Por ejemplo, Carlos Cacho particularizó el uso de **big data en el mundo de los Recursos Humanos**. Aquí tenemos que hacer un punto y aparte, y es que su *startup*, Job & Talent, acaba de conseguir **26 millones de euros**, algo muy relevante para un país como el nuestro. El motivo se denomina **“jobmatching”**: el dato orientado al proceso de captación de los mejores empleados. Cacho nos describió el apasionante crecimiento de esta empresa y cómo está construyendo nuestra **“huella digital” en lo referente a la información profesional**.

Por otro lado, y en un entorno más próximo al de los objetos, tampoco podemos olvidar **uno de los sectores en los que big data será más pródigo en retos**, hablamos de la **fábrica 4.0**, pero sobre todo del **coche conectado**, repleto de *gadgets* que generan 100 Teras/hora de información. David Barrientos nos introdujo en este apasionante mundo y nos dio innumerables pistas sobre los nuevos concesionarios digitales, repletos de pantallas que saciarán nuestra hambre de vehículos atractivos y modernos.

Preguntados los ponentes por titulares de sus mensajes, los resumo: primero, **la tecnología en general y big data en particular deberán trabajar ante todo para las personas**, por ejemplo **en el aspecto médico**, porque **no deben hacernos vivir más... sino mejor**. Y segundo, respecto a todo lo comentado, **es necesario abrir la mente a los cambios...** lo que no significa aceptar cualquier cosa. Porque *big data* nos va a cambiar desde un punto de vista que no conocemos nuestra apreciación sobre **la privacidad y el uso de la información**.

Pero yo destacaré, y así termino, lo siguiente: **todo lo anterior, sin emociones no tiene sentido**. Aunque **muchos ya han tomado conciencia de su importancia** y algunos han empezado incluso a medirlas, otros aún **las ignoran en la toma de decisiones y no pueden permitírselo**. Este efecto se llama **falacia de McNamara: no debemos guiarnos**



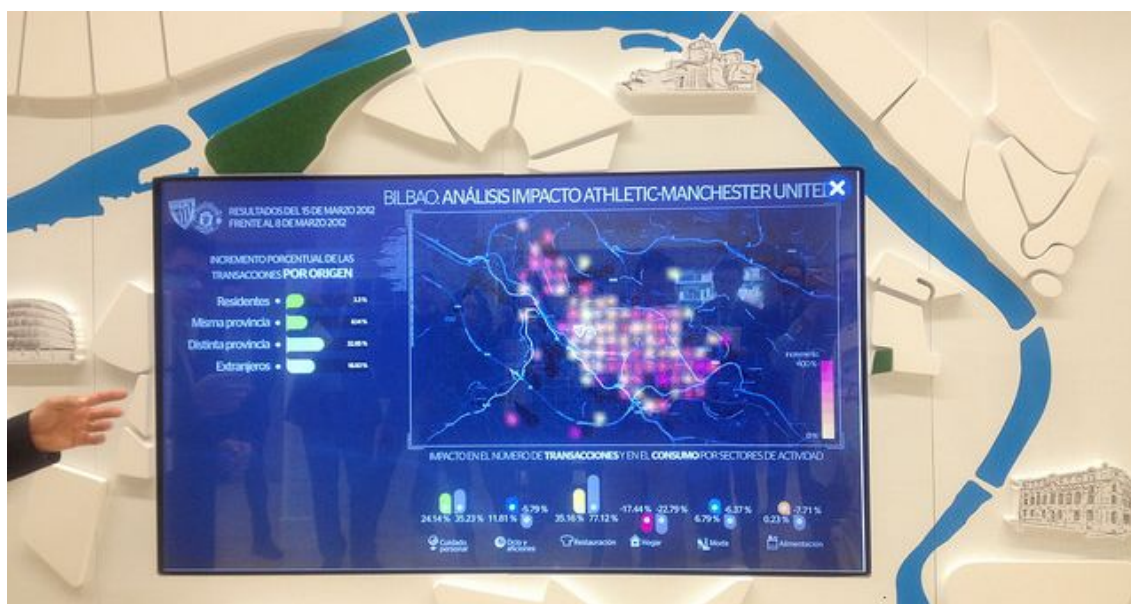
exclusivamente por aspectos cuantitativos. Por suerte, la clave está en el ingenio y para ello tenemos que estrujarnos las meninges e ir un poco más allá y rebuscar en lo desconocido. Buscar detrás de los números..., en las sombras, pues al darles luz, encontramos nuevo conocimiento. Los científicos se pasan el día haciéndolo. Y, si necesitáis más pistas, el olfato de [Clint Eastwood](#) os podrá ayudar un poco.



Sectores pioneros en *big data*

Cómo facilita *big data* la transformación hacia una banca digital

Julián García Barbosa (@juliangbarbosa)



El sector financiero se encuentra inmerso en una encrucijada. Después de casi siete años de la caída de Lehman Brothers ya se atisban los primeros signos de recuperación, aunque **son muchos los desafíos** a los que aún se tiene que enfrentar la banca.

Durante este tiempo, las entidades financieras han tratado de superar los exámenes impuestos por los supervisores, que les exigían unos **severos requerimientos en cuanto a capital, liquidez y endeudamiento**, lo que **ha llevado al sector a un proceso de reestructuración**, concentración y saneamiento sin precedentes.

Una vez superado el escollo de la crisis, los bancos empiezan a levantar cabeza y se encuentran un panorama poco alentador en el que **todo ha cambiado: tanto el terreno de juego como los jugadores**.

Hoy en día una de las preocupaciones fundamentales del sector es la llamada “**banca en la sombra**” (*shadow banking*), que son entidades que, al olor de la crisis y en plena sequía de crédito, empezaron a ofrecer financiación al mercado, al margen de las entidades financieras y sin estar sujetas a ningún tipo de regulación. Dentro de esta modalidad de banca están, por ejemplo, las sociedades de *crowdfunding* (financiación colectiva), *business angels* y fondos de capital riesgo.



Dentro del *shadow banking* se incluyen también las empresas de **fintech** que prestan **servicios financieros a través de la tecnología y fuera de la banca tradicional**, como Google, Amazon, Paypal o Apple, así como infinidad de pequeñas *startups* emergentes que ya operan en determinados segmentos concretos del negocio bancario.

Estos nuevos jugadores tecnológicos están empezando a comerse una parte muy importante de la tarta hasta ahora reservada a la banca tradicional, por lo que actualmente todas **las alertas están activadas** en el sector financiero.

En los últimos años ha sido muy frecuente ver cómo **compañías puramente digitales han irrumpido en distintos mercados y los han transformado, cambiando drásticamente los modelos de negocio** y provocando la **desaparición de todas aquellas empresas tradicionales** que no han sido capaces de reaccionar a tiempo. Por ejemplo, Uber es la mayor compañía del sector del taxi en el mundo y ni siquiera tiene un taxi, Alibaba y Amazon son las mayores empresas del sector *retail* y no tienen tiendas, Facebook es la mayor empresa de contenidos y no genera ninguno propio, y Airbnb es el mayor gestor de alquileres de habitaciones sin habitaciones en propiedad.

A día de hoy **cualquier autónomo o pequeño comerciante podría empezar a aceptar pagos con tarjeta** con solo **convertir su móvil o tableta en un TPV** a través de soluciones que ofrecen compañías de reciente creación como SumUp, o bien **acceder a un servicio de cambio de divisas** mediante la plataforma Kantox. **Cualquier emprendedor podría financiar sus proyectos mediante plataformas de crowdfunding** como Kickstarter. **O las empresas podrían empezar a aceptar monedas virtuales** como **Bitcoin como forma de pago**. Todos estos ejemplos ponen de relieve la **gran cantidad de servicios financieros que confluyen actualmente en el mercado y son ajenos a la banca tradicional**.

Se acerca el invierno para el sector financiero y las entidades no tendrán más remedio que remontar el vuelo para poder enfrentarse a todas las amenazas que acechan en este nuevo escenario.

Sin lugar a dudas, **uno de los mejores activos para los bancos es la gran cantidad de datos** que tienen almacenados de sus clientes. Hay que tener en cuenta que **toda operación, por mínima que sea** (pagos con tarjetas, transferencias, domiciliaciones, ingresos...) **queda registrada**. Debido además a la **nueva normativa de prevención del blanqueo de capitales**, **todos los usuarios de banca han tenido que identificarse** debidamente y esto ha supuesto un incremento notable de dichos datos.

Ahora bien, los datos por sí solos no sirven para nada. **Lo verdaderamente importante es el conocimiento que se puede extraer a partir de dichos datos**, y es aquí donde entra **big data**.

Gracias al uso eficiente de tecnologías de *big data*, así como al procesamiento de un



volumen ingente de datos, de gran variedad y a una gran velocidad, los bancos serán capaces de entender mejor el comportamiento de sus clientes y podrán plantearles ofertas personalizadas en el momento oportuno y, así, proporcionarles la mejor experiencia de usuario posible.

Como casi siempre, la teoría es mucho más fácil que la práctica, ya que implementar soluciones de *big data* en las entidades financieras es de todo menos trivial. Lo más importante para que una estrategia de este tipo resulte efectiva es que la alta dirección apueste rotundamente por ella, y que internamente la impulsen las áreas de Negocio y Marketing.

Big data permitirá a los bancos emprender un proceso de transformación hacia una banca digital que sea capaz de ofrecer a sus clientes acceso sencillo a servicios bancarios personalizados, justamente cuándo y dónde lo precisen (banca contextual).

Además, *big data* podrá ayudar a las entidades financieras a optimizar su gestión de riesgos (tanto de crédito, de liquidez, de mercado, operacional o reputacional), así como la detección y prevención del fraude tal y como explicaba en un [post](#) anterior.

La cuenta atrás ya ha empezado. La banca tradicional tiene aún muchos deberes por hacer. Solo aquellas entidades que consigan adaptarse al nuevo ecosistema digital tendrán cabida en el mercado: “renovarse o morir”.



Big data sanitario: el acelerador del conocimiento y la decisión clínica

Visitación López López (@alrakesi)



17

La cantidad de información existente relacionada con nuestra salud es abrumadora. Los centros de salud, los hospitales, la administración pública e, incluso, nosotros mismos como pacientes acumulamos grandes cantidades de datos en formatos muy diversos: informes en papel, archivos de Office, imágenes, vídeos, recetas, tarjeta sanitaria, etc.

Hasta que se instauró la Historia clínica electrónica cada miembro de la comunidad sanitaria tenía una visión parcial del paciente, lo que dificultaba el diagnóstico y tratamiento. Aunque hoy en día el historial médico de un paciente es compartido dentro del sistema sanitario público autonómico, la situación sigue siendo complicada cuando sales fuera de tu autonomía, o gestionas tu salud en el sector privado.

Es admirable que, a pesar de tener la información poco organizada e integrada, los profesionales de la salud consiguen resultados excelentes en la práctica clínica.

¿Alguien puede imaginarse cómo serían esos resultados si todos los datos asociados a un paciente se pudieran agrupar, consolidar y analizar en un lugar único y accesible? Los profesionales podrían tomar mejores decisiones. ¿Y cómo se podría mejorar la práctica clínica en general, no solo para un paciente, sino para toda una



patología, si se pudieran agrupar, normalizar, y analizar los resultados de todos los pacientes de forma anonimizada? **La lista de posibles aplicaciones es infinita**. Por poner algunos ejemplos, se podrían crear herramientas de soporte a la decisión clínica según datos estadísticos fiables, desarrollar algoritmos predictivos para **evitar en lo posible determinadas enfermedades** o mejorar la eficacia de los medicamentos.

Aunque en este blog se ha escrito ya mucho de *big data* me gustaría recordar a qué nos referimos con este concepto: llamamos *big data* a datos muy grandes, estructurados o no, estáticos o dinámicos, simples o complejos, que pueden ser capturados, almacenados, gestionados y analizados, tanto de forma convencional como utilizando técnicas innovadoras. Como ocurre con la mayoría de los campos de aplicación de ***big data***, **en el caso del sector sanitario es preciso, con mayor motivo aún, que la información agregada y normalizada esté anonimizada**, para cumplir así con el máximo nivel de **seguridad** que exige la actual LOPD (Ley Orgánica de Protección de Datos).

Aunque se trata de un concepto relativamente nuevo, *big data* no ha surgido de la nada, es una evolución del *Business Intelligence* (BI), que utilizan sobre todo los gestores para convertir sus empresas en organizaciones eficaces y eficientes. El BI tradicional captura información de las fuentes disponibles de la organización y, tras la aplicación de algoritmos de análisis, la muestra con el fin de ayudar a la toma de decisiones estratégicas en la empresa. Toda la información analizada y visualizada desde BI proviene de fuentes de datos estructuradas. **La gran novedad de *big data* es el procesamiento de información no estructurada**: lenguaje natural, redes sociales, *wearables*, telemedicina, sensores... Se trata de una **ingente cantidad de información que puede proporcionar datos muy valiosos al sector**.

Todos parecen estar de acuerdo con el gran valor que puede aportar ***big data*** en el mundo de la salud. Los profesionales sanitarios cada vez entienden mejor que **puede suponer un cambio de paradigma en la práctica de la Medicina**, y las empresas farmacéuticas quieren utilizarlo para diseñar medicamentos cada vez más efectivos y con menor coste de investigación. A su vez, las administraciones quieren comprobar la eficacia de los nuevos medicamentos en lo que se denomina "Real World Data". Este nuevo concepto complementa los ensayos clínicos controlados realizados por los laboratorios para determinar la seguridad y eficacia de un nuevo fármaco. Se trata de comprobar "en el mundo real" la eficacia del medicamento una vez comercializado, en un entorno sin restricciones y considerando la amplia casuística de pacientes de una patología y su comportamiento real.

Big data también puede disminuir costes en un sector cuyo **sostenimiento económico empieza a ser cada vez más incierto**. Es evidente que se necesitan fuertes inversiones iniciales para optimizar las técnicas de *big data*, pero también lo es que una vez desarrolladas e implantadas estas nuevas tecnologías supondrán grandes



ahorros a todas las organizaciones sanitarias, públicas y privadas. Según Gartner, en EE. UU. se estiman unos ahorros de 300.000 millones de dólares anuales derivados de la aplicación de *big data* en el entorno sanitario. Además, la consultora estima entre cinco y diez años el tiempo necesario para alcanzar la plena productividad en *big data* sanitario. Por tanto, el que no haya empezado ya a pensar en *big data* no conseguirá con éxito los nuevos retos que se plantearán en los próximos años.

En España ya existen varias iniciativas de *big data*. Cataluña lleva desde 2012 con un [proyecto](#) que verá la luz a finales de este año, en el que se pondrá a disposición de centros de investigación y organismos sanitarios catalanes públicos los datos existentes en los sistemas de información sanitarios de Cataluña.

Otras comunidades autónomas han publicado ya su interés en *big data* sanitario o están estudiando su estrategia de cara a los próximos años, lo que demuestra el grado de concienciación sobre la necesidad de adoptar nuevas tecnologías para optimizar los recursos sanitarios.

Sin embargo, también hay sombras en este **horizonte claro**. Actualmente existen muchas **reservas sobre el fin último que se le puede dar a *big data***. No es lo mismo utilizar datos de pacientes con fines de investigación que proporcionar los datos a empresas que puedan utilizarlos únicamente en beneficio propio. Además, los expertos en protección de datos alertan de que deben existir mecanismos seguros, fiables y consistentes que permitan la total anonimización de los datos de los pacientes e impidan la identificación del sujeto mediante la vuelta atrás del proceso.

En definitiva, el sector de la salud en España ya ha comenzado la andadura y, aunque todavía queda mucho camino por recorrer y muchos obstáculos que vencer, los beneficios que se conseguirán con el *big data* sanitario harán que merezca la pena el esfuerzo.



Big data en Sanidad, clave en la creación de tratamientos eficaces y eficientes

Visitación López López (@alrakesi)



Durante años, los ensayos clínicos aleatorizados (ECA) se han considerado la clave para generar la evidencia científica de la eficacia y la seguridad de un medicamento, la única vía para registrar primero y prescribir después con garantías un principio activo. Sin embargo, recientemente se está debatiendo mucho sobre las capacidades del *Real World Data* (RWD), -en español, los datos del mundo real- y su potencial contribución a la mejora de este campo.

Los ECA son esenciales para verificar la eficacia de un nuevo principio activo o para certificar nuevos usos de medicamentos que ya están en el mercado. Sin embargo, no son válidos para certificar la eficiencia del principio activo. Dicho de otro modo, **los ECA demuestran que un medicamento puede solucionar un problema de salud, pero no son capaces de abordar la amplia casuística** de pacientes reales y los múltiples factores que disminuyen significativamente el número de personas para las que el medicamento es eficaz.

Real world data, que no es más que un estudio observacional basado en datos reales obtenidos de la práctica clínica diaria, no puede sustituir a los datos de seguridad y eficacia generados por los ECA, pero sí puede ayudar a complementarlos pues incorpora información de un gran número de pacientes en un entorno clínico real.



La mayor parte de los pacientes que son atendidos en los servicios sanitarios no participan en los ensayos clínicos. La información derivada de su atención se registra localmente en el peor de los casos o, lo que es más habitual, en la Historia Clínica Electrónica de su Comunidad Autónoma. Gracias al RWD es posible recopilar la información y experiencia de un elevado número de pacientes tratados con el medicamento en estudio y evaluar la eficiencia del principio activo en un entorno real, teniendo en cuenta múltiples variables que no pueden contemplarse en un ensayo clínico.

Tecnológicamente hablando, el RWD es [una de las aplicaciones sanitarias más valiosas del big data](#). La capacidad de procesamiento y análisis de grandes cantidades de datos de fuentes diversas, estructuradas o no, que nos proporciona la tecnología *big data*, hace posible responder a numerosas cuestiones utilizando la información de la práctica clínica real. Mientras **los ensayos clínicos aleatorizados** siguen siendo **la llave para la demostración de la eficacia clínica**, **otros diseños basados en RWD pueden contribuir** de forma efectiva **a mejorar** su validez externa, es decir, **su efectividad y su aplicabilidad a un mayor número de pacientes**.

Las **aplicaciones del *real world data* en el entorno sanitario** son múltiples:

- Desde el punto de vista de las patologías y sus poblaciones de pacientes, **permite analizar las tendencias epidemiológicas y los patrones de tratamientos y enfermedades, así como la adherencia al tratamiento** de los pacientes.
- Permite ajustar el precio de mercado de los nuevos tratamientos, combinando no solo la eficacia comprobada en los ECA, sino también la efectividad contrastada en el entorno real.
- A la hora de desarrollar nuevos productos y terapias, RWD puede **evaluar el uso de los productos actuales en el mercado e identificar los usos no deseados e indicaciones inadecuadas, mejorar los criterios de inclusión/exclusión de los ensayos clínicos** e identificar a los pacientes para el reclutamiento, **o realizar modelos predictivos sobre ensayos virtuales**.
- Finalmente, es importante para **identificar productos y servicios clave mediante el descubrimiento de poblaciones de pacientes desatendidos o con la identificación de subpoblaciones con respuesta superior o inferior al producto**.

El excelente trabajo de la Fundación Gaspar Casal [“Aportación de los Real World Data a la mejora de la práctica clínica y del consumo de recursos de los pacientes”](#), que aunque no es reciente es totalmente actual, nos da una visión de los **enormes beneficios que el RWD puede aportar al entorno sanitario** y de los cambios necesarios para que este nuevo enfoque se instale en la forma de entender la Sanidad.



Entre los beneficios del RWD comparados con los ECA señalados por la Fundación Gaspar Casal destacamos los siguientes:

- Se consiguen **estimaciones de la efectividad más que de la eficacia** en una variedad de escenarios clínicos diferentes.
- **Permite la comparación** con otras alternativas de investigación o estrategias clínicas **para informar de opciones terapéuticas óptimas**, más allá del uso del placebo como comparador.
- Se obtienen **resultados clínicos en una población diversa** que reflejan el rango y la distribución de los pacientes observados en la práctica clínica.
- **Se pueden utilizar estos datos para analizar los costes de los servicios de salud.**
- Se obtiene **información muy valiosa de cómo un producto es aplicado y prescrito en la práctica clínica y cuál es su grado de adherencia.**

Si queremos que el RWD se implante definitivamente como una herramienta de validación clínica queda un importante camino por recorrer. La administración, los profesionales, la legislación y los propios pacientes tienen que cambiar su punto de vista y adoptar una nueva perspectiva:

- **Es imprescindible un cambio cultural.** Todos los agentes implicados deben ser conscientes de la importancia del RWD para estrechar el *gap* de conocimiento existente entre el proveniente de los ensayos clínicos aleatorizados y la gran cantidad de datos, información y también conocimiento, procedente de la práctica clínica real. Para ello es preciso reconocer las necesidades de información que pueden cubrir estos RWD para la toma de decisiones de cada uno de los actores que, directa o indirectamente, participan en los diferentes ámbitos de la Sanidad (política sanitaria, gestión sanitaria, práctica clínica, investigación, docencia...).
- **La tecnología también tiene que aumentar la disponibilidad y accesibilidad a los RWD.** La **interoperabilidad** entre los diferentes sistemas debe aumentar para evitar la fragmentación de la información entre hospitales, comunidades autónomas y diferentes sistemas de provisión de servicios (públicos y privados). Para lograr este objetivo es necesario dotar de **recursos financieros específicos que permitan acortar esta brecha** tecnológica.
- **La legislación debe permitir el uso del RWD con todas las garantías** de privacidad de la información y los aspectos éticos ligados a su utilización y, al mismo tiempo, que atienda a las cuestiones que permitan mejorar la práctica clínica y el uso eficiente de recursos.



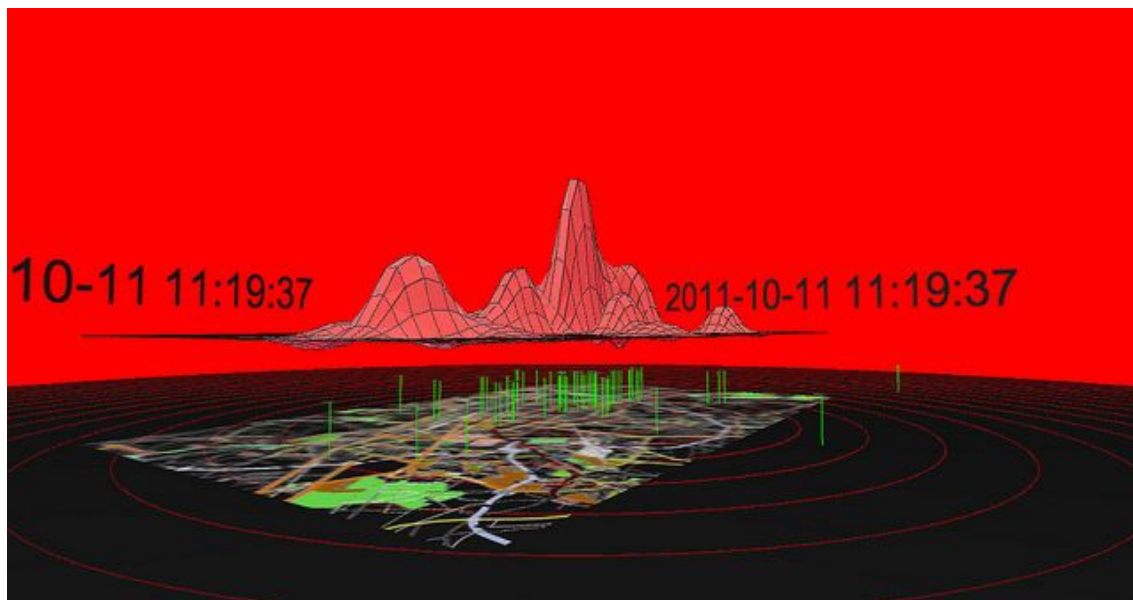
- Es importante **fortalecer el papel del paciente** para que participe en la recolección de los resultados relevantes.
- Se debe incrementar la validez interna de los RWD, mejorando la calidad y exhaustividad de los resultados registrados, **y fomentar la investigación clínica en efectividad comparada a partir de RWD de calidad**, que facilite la mejora de la práctica clínica, evite lagunas de conocimiento y genere conocimiento aplicable a la docencia y educación sanitaria.

En definitiva, ***big data*** se está instaurando en la Sanidad como una **herramienta imprescindible para mejorar no solo la práctica clínica y la investigación médica, sino también como ayuda en la creación de tratamientos eficaces y eficientes** para los pacientes, lo que redundará en una mejor calidad de vida para la persona.



Big data: piedra angular de las smart cities

Julián García Barbosa (@juliangbarbosa)



Contar con redes eléctricas inteligentes, definir las rutas óptimas de recogida de basura en tiempo real, poder anticiparse a los atascos o conocer el sentir de los ciudadanos en cada momento para involucrarlos en las decisiones adoptadas son algunos ejemplos de la “magia” del *big data*.

Según un informe de las Naciones Unidas, en 2050 el 70 por ciento de la población mundial vivirá en entornos urbanos. Pese a ocupar solo el 2 por ciento del territorio del planeta, las ciudades representan entre el 60 y el 80 por ciento del consumo energético mundial y generan el 70 por ciento de las emisiones de gases de efecto invernadero. Todo ello supone que las urbes del futuro tendrán que enfrentarse a grandes retos relativos a la gestión de recursos escasos y a la provisión de servicios tan importantes como la seguridad, el transporte y el abastecimiento de agua y energía, como apuntaba una compañera en “[De ciudades listillas a ciudades inteligentes](#)”.

En este entorno, las TIC se convierten en una herramienta fundamental y *big data* en la piedra angular de las *smart cities*

Como señalaba en “[Big data, el gran trampolín hacia la innovación](#)”, asistimos a una verdadera explosión de datos generados fundamentalmente por las interacciones de las personas en las redes sociales y de los miles de sensores y dispositivos conectados



(Internet de las cosas) y es precisamente en las ciudades donde se concentran la mayoría de la población y de todos estos dispositivos.

Además del gran **volumen** de datos que hay que gestionar en las urbes de hoy en día, estos son de una gran **variedad** (más del 80 por ciento son no estructurados) y es necesario tomar decisiones a una gran **velocidad** (en caso de tener que desviar el tráfico debido a un incendio, por ejemplo). Éstas son precisamente las **3 Vs** que caracterizan todo proyecto de **big data**.

Pero, ¿cómo podrían las ciudades **beneficiarse de las soluciones de big data** para convertirse en verdaderas *smart cities*? Antes de nada, es preciso identificar perfectamente los problemas que se quieren resolver y definir una estrategia única para acometer el proyecto como un todo, implicando a la totalidad de actores involucrados. Es preciso pasar de un modelo de gestión de servicios de forma aislada a tener una visión holística de todo lo que acontece en la ciudad.

Como en cualquier proyecto de *big data*, en el caso de las *smart cities* será preciso capturar, almacenar, procesar y analizar gran cantidad de datos procedentes de fuentes muy diversas para poder transformarlos en conocimiento útil para la toma de decisiones y poder anticiparnos a lo que va a pasar.

Algunas de las áreas susceptibles de mejora gracias al uso de *big data* serían las siguientes:

- 1 Seguridad ciudadana:** Se podría mejorar la eficacia de las actuaciones de los cuerpos de seguridad a través de la correlación de toda la información procedente de los distintos sistemas instalados en la ciudad: desde cámaras de videovigilancia, geolocalización de coches de policía y bomberos, sensores de movilidad o de alertas, detectores de humo y fuego.
- 2 Movilidad urbana:** Mediante la captura y gestión de datos procedentes de cámaras repartidas por toda la ciudad, sensores instalados en autobuses, información meteorológica, datos originados en las redes sociales (como por ejemplo la organización de una manifestación a través de Twitter) se podría conseguir por ejemplo anticiparse a los atascos y tomar decisiones en tiempo real para redirigir la ruta de autobuses o interactuar con la red de semáforos e informar al ciudadano de la situación del tráfico e indicarle las rutas alternativas óptimas.
- 3 Gestión del agua:** A través del análisis de los datos ofrecidos por una red de sensores de presión, PH y turbidez del agua ubicados en los sistemas de abastecimiento y saneamiento, así como cámaras de vigilancia de plantas potabilizadoras, sería posible detectar fugas y controlar la calidad del agua en todo momento.



- 4 **Energía y eficiencia energética:** Gracias a los datos procedentes de contadores inteligentes en las viviendas, datos procedentes de plataformas *open data* y previsiones meteorológicas (que permitieran detectar una inminente ola de calor, por ejemplo), sería posible tener una operación más eficiente de la red eléctrica, de forma que las compañías pudieran ajustar la producción a la demanda en tiempo real (pilar de las *smart grids* o redes eléctricas inteligentes).
- 5 **Residuos urbanos:** Mediante sensores ubicados en los contenedores que envían datos relativos a su capacidad en cada momento, se podrían definir las rutas óptimas de recogida de basura en tiempo real combinando la información de dichos sensores con otros datos procedentes, por ejemplo, de ciudadanos que pudieran estar expresando sus quejas sobre la suciedad de algunas calles a través de redes sociales.
- 6 **Análisis de sentimiento del ciudadano:** Posibilidad de conocer la opinión de los ciudadanos y turistas sobre la ciudad a través del análisis en tiempo real de datos procedentes de distintas redes sociales, web del ayuntamiento, *call centers*... para conocer cuáles son los aspectos prioritarios que están demandando y poder responder a peticiones de forma inmediata.

En todos los entornos de aplicación vemos que una pieza clave son las soluciones **M2M (Internet de las cosas)** -[Santander](#), con más de **12.000 sensores distribuidos por toda la ciudad**, se ha convertido en faro y laboratorio de la inteligencia urbana-, y muy relacionadas estarán también las tecnologías de **cloud computing**, que es una de las vías más interesantes para que los ayuntamientos puedan acceder de forma sencilla y económica a las soluciones de *big data* incluidas en cualquier proyecto de *smart city* (para profundizar en ello os invito leer mi *post* "[Cloud Computing, el mejor camino para acceder al big data](#)").



Big data y la nueva revolución del sector agrícola

Julián García Barbosa (@juliangbarbosa)



27

Según previsiones de la [FAO](#) (Organización de las Naciones Unidas para la Alimentación y la Agricultura), en el año 2050 la población mundial habrá crecido del orden de un 34 por ciento y, para hacer frente a este incremento de la demanda, será necesario duplicar la producción de alimentos.

Por otro lado, nos encontraremos con una disminución de la tierra dedicada al cultivo, ya que muchos de los espacios actuales serán empleados para la creación de biocombustibles. Además, hay que tener en cuenta el progresivo éxodo rural que hará que [la población se concentre cada vez más en las ciudades](#).

El panorama es cuando menos desalentador. **La sociedad se enfrenta al reto de tener que producir el doble de alimentos en una menor superficie cultivable, con menos mano de obra y con graves problemas de escasez de agua.**

Si en la década de los 60 se produjo la “[Revolución verde](#)”, que consiguió incrementar la productividad agrícola mediante la mejora y selección de semillas, evolución de los sistemas de riego, fertilizantes, maquinaria, etc. en este momento **el sector agrícola precisa de una nueva revolución cuyo motor sean las Tecnologías de la Información** y que permita afrontar los retos descritos a través de una “[agricultura inteligente](#)” o “[agricultura de precisión](#)”.



Tradicionalmente los agricultores se han basado en su intuición y experiencia para llevar a cabo las distintas tareas del campo siguiendo un calendario predefinido. Hoy en día, y **gracias al uso de las tecnologías de *big data*, podrán tomar las mejores decisiones en cada momento** a través de la **captura, almacenamiento y análisis de una ingente cantidad de datos en tiempo real procedentes de sensores** que miden el **PH o humedad de la tierra**, calidad del aire o niveles de nitrógeno, imágenes aéreas procedentes de satélites o **drones**, previsión del tiempo por parte de las agencias meteorológicas...

Big data también permitirá mejorar la producción al proporcionar información sobre qué tipo de cultivo es más rentable plantar, en qué fecha exacta y en qué terreno concreto o qué cantidad exacta de agua y fertilizante precisará, lo que **permitirá minimizar los costes de producción, así como los niveles de contaminación**. Y es que **la agricultura del futuro será, ante todo, sostenible con el medio ambiente**.

Distintos tipos de **sensores instalados en tractores, cosechadoras y maquinaria agrícola** en general **informarán al agricultor de posibles averías antes de que ocurran**, lo que permitirá por ejemplo reemplazar una determinada pieza de manera proactiva y así evitar una posible rotura y la inutilización de la máquina en cuestión durante varios días, de lo que se resentiría la **productividad**.

Además de mejorar las cosechas, **las granjas de animales también se verán beneficiadas por *big data***. Mediante el uso de una serie de sensores en los cobertizos se podrá pesar automáticamente al ganado y ajustar la comida en la dosis adecuada en el momento preciso y será posible monitorizar la salud de los animales en tiempo real de forma que aquellos que estén enfermos puedan recibir la medicación necesaria a través de su comida.

Las enfermedades, plagas y las propias condiciones meteorológicas convierten la agricultura y la ganadería en negocios de altísimo riesgo. Todos los años se producen casos de cosechas perdidas totalmente por sequías, nevadas o lluvias torrenciales, así como el cierre de granjas debido a enfermedades como los famosos casos de las “vacas locas”, la peste porcina o la gripe aviar. **Mediante *big data* y el uso de analítica predictiva**, agricultores y ganaderos podrán anticiparse al futuro con la toma de decisiones oportunas para, al menos, poder **reducir el impacto ante posibles catástrofes**.

Pero la agricultura del futuro no solo será aquella que pueda mejorar los niveles de producción con el menor uso posible de terreno, agua, fertilizantes y mano de obra, sino que **además tendrá que permitir optimizar los márgenes de venta de sus cosechas**.

Y es que la gran volatilidad de precios en este sector es una de las mayores pesadillas



de los agricultores, porque las oscilaciones que se producen son muy bruscas y repentinas: en solo dos meses se ha llegado a duplicar el precio internacional del arroz. Gracias a la tecnología de *big data*, el agricultor podrá anticiparse a variaciones importantes en el precio de determinados productos mediante el uso de modelos matemáticos y recibir recomendaciones en tiempo real para comprar y vender al mejor precio posible en cada momento.

Tenemos, por tanto, todas las esperanzas puestas en que la agricultura inteligente sea en breve una realidad para afrontar con éxito el reto de obtener los niveles de producción adecuados con el menor impacto posible en el medio ambiente y poder así **asegurar el alimento para las sociedades del futuro.**



Big data en las aulas: del “café para todos” a una “educación a la carta”

Julián García Barbosa (@juliangbarbosa)



Supongamos que un ciudadano del siglo XIX emprendiera un viaje a través del tiempo y aterrizara en nuestra época actual. Es fácil imaginar que quedaría absolutamente fascinado con los grandes avances tecnológicos que han tenido lugar en todos los sectores. Sin embargo, si visitara cualquier escuela o universidad le podría parecer que todo se mantiene igual.

Posiblemente percibiría algunos cambios menores y es que, después de las múltiples **reformas educativas** que se han llevado a cabo en todos los países, se han mejorado algunos aspectos: el ambiente en la clase es mucho más distendido, el profesor fomenta la colaboración entre los alumnos, hay equipamientos audiovisuales en las aulas, los alumnos disponen de portátiles o tabletas y pueden consultar Internet para la realización de sus trabajos... Pero todas estas **mejoras** no dejan de ser “**cosméticas**” y no han llegado nunca a modificar el núcleo del sistema educativo.

Me explico: **el aprendizaje es una experiencia individual** y específica para cada persona. Es el resultado de procesos cognitivos propios mediante los cuales asimilamos e interiorizamos nuevas informaciones que percibimos a través de los sentidos y que codificamos y almacenamos en nuestra memoria. Pero si cada alumno es único y procesa la información de forma distinta, ¿por qué a día de hoy todos los libros de textos, las exposiciones de los maestros en las clases, los deberes y evaluaciones son exactamente los mismos para todos los estudiantes, sin hacer ningún



tipo de distinción?

En todos los sectores el consumidor puede elegir un producto o servicio entre un sinnúmero de opciones, pero **en educación seguimos teniendo el “café para todos”**.

En la antigüedad llegó a ser posible una educación personalizada, pero solo la iglesia, los nobles y **ricos** en general **se podían permitir el lujo** de tener un profesor particular. Desde el proceso de democratización de la enseñanza que se produjo en el siglo XIX y hasta la fecha se ha tenido que prescindir de las bondades de la **“educación a la carta”** en pro de ofrecer un sistema educativo gratuito a toda la sociedad.

Actualmente los Cursos masivos abiertos *on line* (**MOOC**) están ganando **cada vez más popularidad**, las editoriales publican cada vez más libros de texto escolares y universitarios en formato **eBook**, los alumnos y profesores usan habitualmente *smartphones* y *tablets* para acceder a contenidos educativos, las redes de datos móviles de alta velocidad permiten al alumno estar conectado en cualquier lugar... Todo esto hace que el terreno de juego esté perfectamente preparado para emprender una verdadera **transformación de la educación**, y la chispa necesaria para que este proceso comience pasa por implementar correctamente las tecnologías de **big data**.

Hoy en día, gracias al uso de las nuevas plataformas de educación digital, sería posible capturar, almacenar y procesar todos los datos relativos a cada alumno en tiempo real a lo largo de todo su proceso de aprendizaje. Podríamos así saber cuánto tiempo ha tardado en responder una determinada pregunta de un examen, cuántas y qué cuestiones se ha saltado sin responder, si ha necesitado revisar una determinada lección varias veces, si ha tenido que volver a repasar determinados conceptos de lecciones anteriores para poder avanzar o qué consejos por parte de los profesores han funcionado mejor en cada alumno.

Con una **monitorización de los alumnos** se podrían identificar las áreas específicas en las que cada uno de ellos necesita ayuda y personalizar las lecciones y contenidos educativos para optimizar la comprensión de las distintas materias.

Gracias a **big data** los profesores podrán definir una **ruta personalizada** para los estudiantes durante todo el proceso educativo. Se trata de que cada alumno pueda convertirse en la “mejor versión de sí mismo”, que es el fin último de la educación. En este nuevo escenario **el profesor** pasa a ser fundamentalmente un **coach** de sus discípulos para intentar descubrir sus debilidades y fortalezas, y guiarlos en el día a día, con motivación, para superar los obstáculos y fomentar sus puntos fuertes.

Hasta ahora el único *feedback* existente en el sector educativo era en una dirección: desde los profesores/colegios a los alumnos/padres según las calificaciones de exámenes, comportamiento y participación en clase o la realización de deberes, como



únicas fuentes de información. Siempre **nos hemos centrado en evaluar los resultados de la formación, olvidándonos de medir el proceso de aprendizaje** en sí mismo. Pero *big data* ha llegado a las aulas para dar un giro radical a esta situación con un **feedback en tiempo real** tanto para los estudiantes, para que puedan ser más eficientes en sus estudios, como para los profesores, para que sean conscientes de qué técnicas y métodos funcionan en cada alumno, y para que las editoriales puedan mejorar y corregir los libros de texto digitales de forma dinámica.

Y todo esto ya es posible porque *big data* es una realidad, aunque aún queda **lo más difícil** por hacer, y es **que los todos los agentes involucrados** (gobiernos, universidades, centros educativos y empresas de servicios tecnológicos) **se alineen y emprendan el proyecto de transformación definitivo** que el sector de la educación requiere para, de esta forma, obtener **una sociedad mejor preparada, más creativa, con una menor tasa de desempleo y, en definitiva, más competitiva.**



Big data en el deporte: ¿dopaje tecnológico o mejor toma de decisiones?

Roberto García Esteban (@72rge)



33

Es un tema de actualidad, hace poco tiempo podíamos leer que “[Ellas entrenan con drones y espinilleras inteligentes](#)”. El uso de **big data** es cada vez más popular en las [finanzas](#), la [salud](#), el marketing... pero también **en el deporte profesional**. Tanto que el 80 por ciento de los equipos de la NBA, la Liga de Baloncesto americana, cuenta con un experto en *big data* en sus filas.

Existen medios cada vez más sofisticados para recoger y posteriormente analizar enormes cantidades de datos relacionados con múltiples aspectos de cualquier deporte: cámaras, sensores y *wearables* guardan cada detalle de la actuación de un deportista. Con esos datos los preparadores y entrenadores pueden pulir defectos en la técnica, mejorar su colocación en el terreno de juego, ajustar la cantidad de calorías que debe ingerir el deportista en función del esfuerzo y desgaste que hace de ellas o los niveles de entrenamiento que puede soportar sin riesgo de lesiones. Todo ello **con el objetivo de lograr el mejor rendimiento posible**.

En Europa, un equipo profesional **pionero en el uso del big data ha sido el Arsenal**, club de la *Premier league* de fútbol inglesa. Recientemente ha instalado en su estadio ocho cámaras que siguen en todo momento a sus jugadores y recogen la friolera de **1,4 millones de datos por partido**. Se centran especialmente en recoger los momentos en los que los futbolistas no están en contacto con el balón, que es la mayor parte del tiempo durante un partido. Analizan cómo se desmarcan, cómo tratan de obstaculizar



el juego del contrario o cómo se colocan cuando el balón sale fuera. En la misma línea, el FC Barcelona también ha implantado de la mano de Telefonica un **sistema** que le permite conocer más en detalle los patrones de juego de los equipos rivales y mejorar los movimientos de sus jugadores.

Otro caso muy significativo es el de los **Golden State Warriors**, equipo de baloncesto de la NBA, ubicado por cierto en San Francisco, muy cerca de Silicon Valley. En 2013 era el equipo que más lanzamientos de tres puntos en contra recibía, así que instaló un **sofisticado sistema de cámaras** llamado “*SportVu*” que básicamente divide la pista en una cuadrícula tridimensional y graba veinte imágenes por segundo de cada jugador para así identificar los **fallos de colocación**. Ha debido dar sus frutos porque **este año han sido los ganadores de la NBA**.

Los Golden State Warriors también son pioneros en aplicar los **wearables en sus entrenamientos**. La equipación de los jugadores contiene nada menos que dieciséis sensores que recogen **datos del ritmo cardiaco, la respiración y la actividad de los principales grupos musculares**. Un pequeño dispositivo en el pantalón recoge toda esta información y la envía a un *Smartphone* desde el que puede analizarse la actividad del jugador en tiempo real.

El uso de **wearables** se está extendiendo rápidamente en el deporte profesional. Pequeños sensores en los vendajes o en la ropa deportiva recogen datos, como la velocidad del deportista o su ritmo cardiaco, que luego **permiten** a científicos, nutricionistas y preparadores físicos **diseñar programas de entrenamiento personalizados** y llegar a conocer cómo el entrenamiento está afectando a su salud o si se está elevando peligrosamente el riesgo de lesión. Hay equipos que **analizan incluso si sus deportistas duermen bien o mal** mediante información recogida por dispositivos que llevan mientras descansan y que luego se relacionan con su actuación durante la actividad deportiva para, por ejemplo, llegar a regular exactamente la temperatura y humedad de las habitaciones de los atletas para que su descanso sea el óptimo y puedan lograr el máximo rendimiento.

En deportes de mucho contacto, como el fútbol americano o el rugby, los jugadores también llevan pequeños sensores en su ropa que miden la dureza e intensidad de los impactos, de manera que tras comparar esas mediciones con una base de datos histórica son **capaces de prever si un jugador está en grave riesgo de lesionarse**.

Hay quien piensa que la aplicación de las modernas técnicas de datos al deporte es una especie de “**dopaje tecnológico**” que aporta a unos deportistas sí y a otros no ventajas provenientes de fuera de la pista o del terreno de juego. Pero **igual que los deportistas de élite se rodean de nutricionistas, psicólogos, fisioterapeutas etc. de los que tratan de obtener una ventaja competitiva frente a sus rivales**, el siguiente paso parece que será contratar a los mejores analistas de datos.

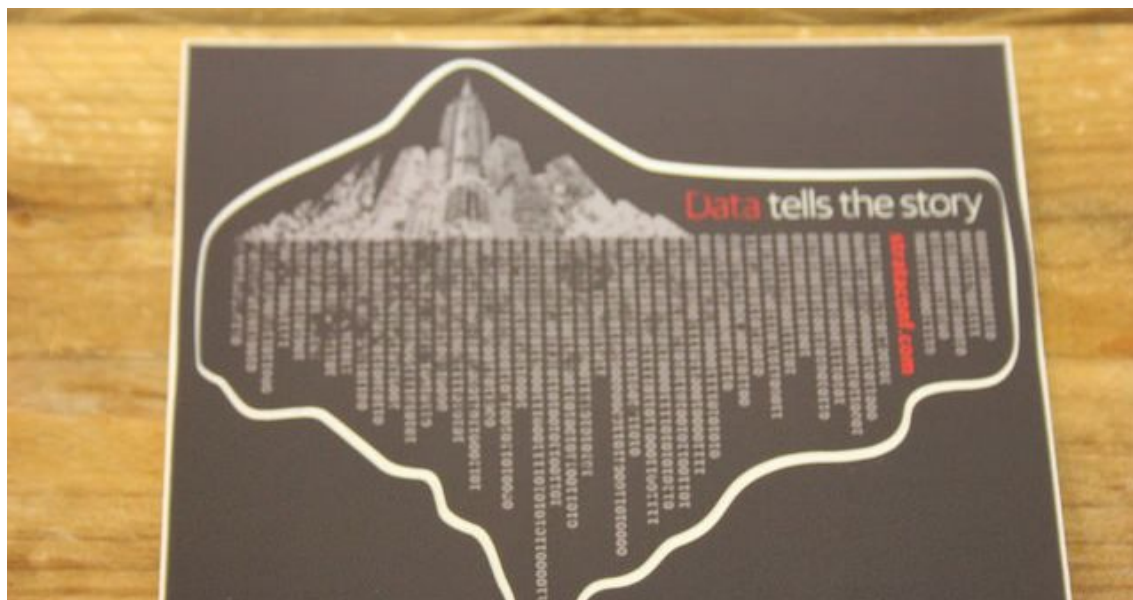


Al final es el deportista el que tiene que lanzar el tiro a canasta o el penalti decisivo, y los análisis de los datos que se hayan hecho sobre él no son capaces de golpear o lanzar un balón. ***Big data* en el deporte de élite tiene como objetivo ayudar a entrenadores y dirigentes a tomar las mejores decisiones** que a la postre servirán para que los deportistas de élite obtengan el máximo rendimiento y ofrezcan un mejor espectáculo a los aficionados.



El periodismo de datos como creador de historias

Mercedes Núñez (@Mertrak)



En el mes de junio de 2015 se celebraba el congreso [GEN Summit](#) sobre el momento presente y futuro del periodismo y, en el encuentro, se abordaron interesantes aspectos como que “el 80 por ciento de las redacciones tendrá un sistema unificado de contenidos (CMS) en los próximos cinco años porque cada vez más los periodistas tienen que producir vídeos, infografías animadas, visualizaciones de datos, directos, *podcasts*...” y se destacaba el papel del periodismo de datos para crear grandes historias.

En este sentido, recomiendo echar un vistazo al [vídeo](#) de las “III Jornadas sobre periodismo de datos y datos abiertos” de principios de junio en Espacio Fundación Telefónica.

Allí también se dijo que había que **convertir la relación entre periodista y programador en una historia de amor** porque es una integración necesaria que ya está dando grandes frutos. Cada vez es más aconsejable que el periodista entienda de programación y que el programador sepa leer código con visión periodística. De hecho, el **periodista de datos** aparece como **una de esas nuevas profesiones tan demandadas**. El mundo académico arrastra su habitual lentitud pero es un tema que cada día cobra mayor fuerza. El [Máster en Periodismo de investigación, datos y visualización](#) de Unidad editorial, que en octubre comenzó su IV edición, es un



ejemplo.

[Mar Cabra](#), profesora de dicho Master y miembro del Consorcio Internacional de Periodistas de Investigación (ICIJ) destacó en las Jornadas cómo **este nuevo periodismo se está abriendo paso en los medios españoles** en una estrategia de guerrilla y destacó el caso de la unidad de datos de El Confidencial, que arrancó con un becario que había realizado el Máster. Hoy son dos periodistas y un programador que pueden aportar valor a cualquier sección del periódico.

Dos investigaciones en las que ha participado *este medio* han sido galardonadas con el prestigioso *Data Journalism Award*: ‘[SwissLeaks](#)’ y ‘[LuxLeaks](#)’, ambas coordinadas por el ICIJ y publicadas en colaboración con más de cincuenta medios en todo el mundo. Y es que las macroinvestigaciones de este tipo requieren de trabajo en equipo y del tiempo necesario para hacerlo bien, que incluye el que se dedica al “reanálisis”, que es como un “chaleco antibalas”, se dijo en el encuentro.

La financiación, la independencia, la dificultad para acceder a los datos (marcos legales de *open data* y transparencia), la variedad de fuentes, de formatos, la depuración necesaria... son **escollos** habituales, pero el resultado merece la pena.

El periodismo de datos permite conseguir nuevas historias. Un gran referente es [ProPublica](#), la agencia de noticias independiente y sin ánimo de lucro, con sede en Manhattan, que **crea aplicaciones de noticias con datos**. No trabajan con la actualidad, sino **con historias relevantes que crean responsabilidad y contribuyen al cambio necesario para un mundo mejor**. Un ejemplo es **“[The opportunity gap](#)”** sobre la brecha de oportunidades para un acceso igualitario a la educación en Estados Unidos.

En la era digital **se puede hacer periodismo con menos palabras y también sin ellas**. A veces **los datos revelan historias en sí mismos** y ahí reside la clave. Cada día hay retos distintos y quizá sea el momento de empezar a mirar el Excel con otros ojos (como punto de partida), sin olvidar a Kapuscinski.



Datos: de las estadísticas y las barras a historias para la sociedad

Mercedes Núñez (@Mertrak)



En las “III Jornadas sobre periodismo de datos y datos abiertos”, celebradas en junio de 2015, [Salvador Pérez](#), de Telefónica I+D, nos mostró cómo había cambiado la forma en que los datos “hablan” y que en Telefónica se había evolucionado de mostrar gráficos, estadísticas, barras... a contar historias para la sociedad a través de ellos, y mencionó algunos proyectos que, por su atractivo, quiero destacar:

- En 2012 nace [Smart steps](#), una solución que utiliza datos móviles anónimos y agregados, para **ayudar a las organizaciones a tomar mejores decisiones en base al comportamiento que dichos datos revelan**. Por ejemplo, el volumen y la naturaleza de las personas que transitan por una determinada zona resulta determinante a la hora de abrir un establecimiento. Las estadísticas también pueden contribuir a **optimizar la ruta de autobuses** o a determinar el emplazamiento idóneo de servicios públicos como hospitales o bomberos, como veíamos en el [post](#) “El *big data* mejora el Ebitda”.

- Otro ejemplo sería el uso que se hizo de **la red móvil para monitorizar la efectividad de las medidas del gobierno mejicano para limitar los movimientos de la población por el peligro de contagio de la pandemia de la gripe A**. Se pudo así contrastar su validez, algo vital en estos casos, pero además Telefónica cuenta con el modelo que permitiría simular nuevas medidas y predecir lo que ocurriría.



- El informe de Telefónica I+D y RocaSalvatella “**Big data y turismo: nuevos indicadores para la gestión turística**” tiene una enorme importancia para este sector, de tanto peso en nuestro país, pues de él se pueden extraer valiosas conclusiones y recomendaciones. **Xavier Capellades** nos contó en **Innodata** que, hasta ahora, las decisiones en este sentido se basaban en encuestas pero *big data* introduce nuevas variables e indicadores contrastados, como los horarios de los turistas, los lugares más visitados o el dinero que gastan. Conocer, por ejemplo, que el lugar más visitado de Barcelona es el puerto, y que el 30 por ciento de los turistas que llegan a él no se mueve de allí, debe permitir establecer un plan de acción para capturar esta oportunidad.
 - Aunque seáis capaces de vislumbrar ya las oportunidades que brinda *big data*, puede que esto os deje con la boca abierta: **un algoritmo puede anticipar si se cometerán delitos** en una zona; ya se ha creado un “**mapa de la criminalidad en Londres**” con una precisión del 70 por ciento.
 - Y, por último, **los datos en el fútbol. Es posible analizar la forma de jugar**: número de pases, distancia media que recorren, su relación con el equipo, y hasta el tipo de dinámica que crean sus pases: los sencillos, redes más sencillas, **y se ha demostrado que en el éxito de un equipo influyen los pases complejos**, con más posibilidades para marcar.
- ¿No es apasionante? ¡Ojo al *data*!



Autores



Julián García Barbosa

Licenciado en Ciencias Físicas (especialidad Electrónica) por la Universidad de Sevilla y MBA por la Universidad Politécnica de Madrid, actualmente trabajo en Telefónica en el área de Entidades Financieras. Apasionado de las Nuevas Tecnologías, sobre todo desde el punto de vista de cómo pueden mejorar la calidad de vida de las personas.



De Julián García Barbosa en este libro:

*Cómo facilita **big data** la transformación hacia una banca digital*

*Big data: piedra angular de las **smart cities***

Big data y la nueva revolución del sector agrícola

Big data en las aulas: del “café para todos” a una “educación a la carta”

41



Roberto García Esteban

Ingeniero de Telecomunicaciones por la UPM, diplomado en Administración y Dirección de Empresas en la UNED y MBA Executive por el Instituto de Empresa. Actualmente trabajo como responsable de la Oferta Global para pymes en Telefónica Digital. Soy muy futbolero (por supuesto, del Real Madrid), si bien me gusta seguir y practicar todo tipo de deportes (es fácil que me veas en una carrera popular por las calles de Madrid). Aunque mi verdadera pasión es jugar con mis dos hijos pequeños.



De Roberto García Esteban en este libro:

Big data en el deporte: ¿dopaje tecnológico o mejor toma de decisiones?





Félix Hdez. de Rojas

Vallisoletano. Ingeniero Superior de Teleco y MBA por ESADE. Emprendedor, visionario y apasionado por las TICs. Creo en la transformación y en el gran valor que aporta la hiperconectividad y la IoT a nuestra sociedad de la información. Ya con muchos años de experiencia y cada vez con más ganas de aprender de todo un poco. Me siento un niño recién parido. Dispuesto a vivir nuevas aventuras: El marketing y los contenidos digitales... Quién quiera saber más que me pregunte. O que busque mi nombre en Google. ¡Ah!, también escribo: visitad mi eloterodelalechuza.com o mis novelas, Spanish Texas y Tempus Fugit Est.



De Félix Hdez. de Rojas en este libro:

Big data, una marea imparable

Big data e IoT: un mar de oportunidades para las grandes empresas

Datos al servicio de la persona



Visitación López López

Licenciada en Ciencias Químicas por la Universidad Autónoma de Madrid, y con Ingeniería Técnica de Informática de Gestión por la UNED casi acabada, en la actualidad me dedico a marketing de producto eHealth en Telefónica. Mi tiempo libre lo dedico al estudio de la antropología, la genética, la literatura, la cosmología... y a mi familia.



De Visitación López López en este libro:

Big data sanitario: el acelerador del conocimiento y la decisión clínica

Big data en Sanidad, clave en la creación de tratamientos eficaces y eficientes





Mercedes Núñez

Licenciada en Periodismo y Máster por la Universidad Complutense de Madrid. Responsable del blog A un clic de las TIC, en el que escribo habitualmente; colaboro también en Con tu negocio y Think Big. Mi especialización en periodismo tecnológico empezó en ABC Informática - el primer semanario español TIC. Llevo quince años en Telefónica haciendo comunicación corporativa (relación con los medios, web, revista PULSO -impresa, digital y para tablet-, redes sociales...) Con anterioridad trabajé en distintos medios de comunicación (prensa, radio y televisión). Mi mayor afición es la equitación.



De Mercedes Núñez en este libro:

[El periodismo de datos como creador de historias](#)

[Datos: de las estadísticas y las barras a historias para la sociedad](#)



Comparte este ebook:



Reconocimiento-No Comercial 2.5 España (CC BY-NC 2.5 ES)

Textos: A un clic de las TIC

Imágenes: .Escola Santa Anna, r2hox, Intel Free Press, Jose A. del Moral, Andy Leppard, kev-shine, xdx_d_vs_xdx_d, CIAT, Michael Coghlan, Keith Allison, Medialab Prado, Kris Krug.

Madrid, enero de 2016

A un clic de las TIC

www.aunclicdelastic.com



Telefonica