



CHAD HAGEN

ANÁLISIS

Una explosión de información

Por STEVE LOHR

Bienvenidos a la era de los grandes datos. La expresión "grandes datos" es una forma de referirse a las tendencias tecnológicas avanzadas que abren la puerta a un nuevo enfoque para entender el mundo y la toma de decisiones. Los datos crecen todo el tiempo, a un ritmo del 50% anual, o lo que es lo mismo, más que duplicándose cada dos años, según calcula la empresa de investigación tecnológica IDC.

Los analistas ayudan a las empresas a encontrar sentido a la explosión de datos —el tráfico en Internet y los comentarios en las redes sociales, así como el

software y los sensores que controlan los envíos, los proveedores y los clientes — para orientar las decisiones, reducir los costes y aumentar las ventas.

Un informe elaborado el año pasado por McKinsey Global Institute, la sección de investigación de la empresa consultora del mismo nombre, pronosticaba que Estados Unidos necesita entre 140.000 y 190.000 trabajadores más con experiencia "analítica profunda" y 1,5 millones más de gestores versados en datos. En las economías desarrolladas de Europa, calcula el informe, los administradores gubernamentales podrían ahorrar más de 149.000

millones de dólares en mejoras de la eficacia operativa.

En campos tan diversos como la ciencia, los deportes, la publicidad y la sanidad pública ha habido un cambio de rumbo hacia los descubrimientos y la toma de decisiones basados en los datos. "Es una revolución", afirma Gary King, director del Instituto de Ciencias Sociales Cuantitativas de la Universidad de Harvard, "aunque en realidad no hemos hecho más que empezar".

Los nuevos megarricos de Silicon Valley, que hoy están en Google y Facebook, dominan el

Pasa a la página 4

Una explosión de información

Viene de la página 1

arte de la vinculación de los datos de la Red —búsquedas, publicaciones y mensajes— con la publicidad en Internet. Un informe del Foro Económico Mundial reunido el mes pasado en Davos, Suiza, declaraba que la información es una nueva clase de activo económico, como las divisas o el oro.

No es solo que haya más corrientes de datos, sino que son completamente nuevas. Por ejemplo, ahora hay incontables sensores digitales en todo el mundo en los equipos industriales, los coches, los medidores eléctricos y los cajones de embalaje. Pueden medir y transmitir la ubicación, el movimiento, la vibración, la temperatura, la humedad e incluso los cambios químicos del aire.

Las herramientas informáticas para recoger información del vasto tesoro de Internet están ganando terreno rápidamente. En primera línea se encuentran las técnicas de la inteligencia artificial, como el procesamiento del lenguaje natural, el reconocimiento de patrones y el aprendizaje por parte de las máquinas.

Estas tecnologías pueden aplicarse en muchos campos. El negocio de búsqueda y publicidad de Google y sus coches robóticos experimentales usan

Según Davos, los datos son un nuevo activo económico.

la inteligencia artificial.

La abundancia de datos nuevos, a su vez, acelera los avances en la informática. Siri, la aplicación de Apple para iPhone que habla y responde a preguntas, es un buen ejemplo. Con la ayuda de personas que proporcionan millones de preguntas, Siri se está convirtiendo en un ayudante personal cada vez más hábil.

Erik Brynjolfsson, economista del Instituto Tecnológico de Massachusetts, dice que en los negocios, la economía y otros ámbitos, las decisiones se basarán cada vez más en los datos y los análisis, en vez de en la experiencia y la intuición. “Podemos empezar a ser mucho más científicos”, señala.

Los análisis masivos de datos también se han convertido en algo habitual en los deportes, entre ellos el fútbol y el béisbol, según se narra en el libro (y en la película del mismo nombre nominada al Oscar) *Moneyball*.

Los minoristas analizan los datos de ventas, de precios, económicos, demográficos y del tiempo para adaptar las selecciones de los productos a las tiendas concretas y fijar las rebajas de los precios. Las empresas de transporte exploran los datos sobre las horas de envío de los camiones y los patrones del tráfico para ajustar la elección de las rutas.

Las agencias de contactos de Internet filtran constantemente sus listados electrónicos de características personales, reacciones y comunicaciones

para mejorar los algoritmos que sirven para emparejar a las personas.

Los departamentos de policía usan el mapeo informatizado y el estudio de variables como los patrones históricos de detenciones, los días de paga, los acontecimientos deportivos, la lluvia y las vacaciones para intentar predecir los probables “puntos calientes” de criminalidad y desplegar allí a sus agentes de antemano.

El poder predictivo de los grandes datos se muestra prometedor en ámbitos como la sanidad pública y las previsiones económicas. Los investigadores han encontrado un repunte en las peticiones de búsqueda de Google para términos como “síntomas de la gripe” y “tratamientos para la gripe” un par de semanas antes de que haya un aumento del número de pacientes con gripe que acuden a los servicios de urgencias.

Global Pulse, una iniciativa de Naciones Unidas, quiere que la información contribuya al desarrollo mundial. Este grupo analizará las redes sociales y los mensajes de texto para ayudar a prever las pérdidas de puestos de trabajo, las reducciones del gasto o los brotes de enfermedades en una zona.

Los grandes datos están transformando el estudio del modo en que funcionan las redes sociales, lo que conlleva la exploración de enormes conjuntos de datos digitales sobre el comportamiento colectivo en Internet.

Los investigadores pueden ver patrones de influencia y repuntes en la comunicación acerca de un asunto siguiendo los temas de actualidad en Twitter, por ejemplo. “Busco puntos calientes en los datos, un estallido de actividad que tengo que entender”, dice Jon Kleinberg, catedrático de la Universidad de Cornell en Ithaca, Nueva York.

Pero esta explosión de información tiene sus peligros, desde luego. Los defensores de la privacidad advierten de que puede convertirse en el Gran Hermano, con un disfraz corporativo. Y los estadísticos y los informáticos señalan que, con unos conjuntos enormes de datos y unas mediciones muy refinadas, hay un mayor riesgo de “descubrimientos falsos”.

Los grandes datos también pueden proporcionar materia prima para los engaños estadísticos y las incursiones sesgadas en el descubrimiento de hechos. Ese es, según Rebecca Goldin, matemática de la Universidad George Mason en Virginia, uno de sus “usos más perniciosos”.

Los modelos de los grandes datos son simplificaciones. Son útiles para comprender, pero tienen limitaciones. Un modelo puede encontrar una correlación y extraer una conclusión estadística que sea injusta o discriminatoria, y que afecte a los productos, los préstamos bancarios y el seguro sanitario que se le ofrecen a alguien.

Pero parece que no hay vuelta atrás. “La cultura ha cambiado”, asegura Andrew Gelman, estadístico de la Universidad de Columbia, en Nueva York. “Actualmente, existe la idea de que los números y las estadísticas son interesantes y divertidos: es decir, están de moda”.