

“Big Data” y el elefante de peluche

Ciento cincuenta contactos o amigos. Es el número fijado, hace ahora veinte años, por el antropólogo Robin Dunbar como la cifra límite de contactos con los que una persona puede relacionarse de forma estable sin llegar a la saturación. Es lo que se conoce, en el mundo de las relaciones personales, como el número Dunbar.

POR SANTI RUANA

Esta cifra, ha quedado hoy en día pulverizada con las nuevas herramientas que nos ofrecen las redes sociales. Todos sabemos que el panorama tecnológico está evolucionando rápidamente. Por ejemplo Facebook, que ha conseguido organizar todo un ejército alrededor de su comunidad de desarrolladores, consiguió que su portal fuera traducido al francés en tan sólo un día.

Un error muy común de las compañías es tratar a los clientes como una masa global en lugar de una suma de individualidades. Por ejemplo, las compañías energéticas suelen hablar de contadores inteligentes en lugar de consumidores. Al igual que un avión a reacción deja a su paso una estela en el cielo, nosotros, los consumidores, también dejamos un rastro que revela nuestros gustos, opiniones y elecciones. Es lo que se denomina nuestra huella digital o “digital footprint”.



Esta información sobre el comportamiento de los clientes siempre ha sido relevante para las compañías para poder ajustar mejor su oferta, segmentar de forma más granular sus clientes, realizar promociones más efectivas, etc. Lo que está cambiando es la forma de obtener y analizar esta información. Durante muchos años, la única forma de obtener datos era mediante encuestas. Con la ayuda de las TIC, aparecieron las primeras oleadas de herramientas para el análisis de información, pero siempre confinadas a sistemas y datos internos disponibles dentro la organización. Amazon, por ejemplo, captura información de los usuarios que visitan su web creando perfiles complejos de sus hábitos y preferencias. Las tarjetas de fidelización que utilizan los más de diez millones de clientes de TESCO, es otro ejemplo de como aprovechar la información generada por los consumidores.

Los efectos combinados de las nuevas tecnologías disponibles en internet, la disponibilidad de gran capacidad de computación, y la omnipresencia de rápidas comunicaciones, son punto de inflexión para la aparición de la nueva era de análisis de datos que llamamos “Big Data”. Los mayores beneficios los obtendrán las organizaciones que combinen diversos conjuntos de datos, internos y externos, siendo los últimos procedentes de redes sociales o datos suministrados por otras compañías.

Predecir los comportamientos

Las TIC históricamente han tenido un gran impacto para la innovación y mejora de competitividad y productividad. Para analizar datos no estructurados, originados en diversas fuentes, e intrínsecamente masivos, será necesario integrar almacena-

miento, computación y aplicaciones específicas, de una forma distinta de como se implementan las arquitecturas para aplicaciones tradicionales. Una de las mejores aproximaciones para abordar este tipo de datos es Hadoop. Doug Cutting, lo inventó para Yahoo, y le puso el nombre del elefante de peluche, bastante feo por cierto, que pertenecía a su hijo. Hadoop distribuye los datos a través de extensos “clusters” de hardware “commodity”.

Esta aproximación es muy diferente a la tradicional de “Data warehousing”, donde la información está clasificada en filas y columnas muy ordenadas, en máquinas dedicadas, donde los informes tardan horas, si no días en generarse. ¿Qué os parecería si pudiéramos analizar cada transacción, capturar los comportamientos de todos nuestros clientes, y no tener que esperar meses para obtener esos datos? Ford Motor, PepsiCo, y Southwest Airlines analizan los mensajes que sobre ellos escriben los consumidores en páginas como Facebook y Twitter, así pueden evaluar de forma inmediata las campañas de marketing que están activas.

El extraordinario valor de Big Data impulsará la creación de nuevos sistemas y herramientas para el reconocimiento de patrones que, potencialmente, podrán aplicarse a todos los sectores. En el caso particular de la administración pública será posible interactuar con los ciudadanos de una forma muy distinta. Pensemos, por ejemplo, en encuestas a tiempo real a través de dispositivos móviles, cálculo de índices de precios de consumo con mayor granularidad de productos y presentados diariamente en lugar de mensualmente, o permitir al público el acceso a todo tipo de información gubernamental como

implementó en 2010 el gobierno federal de EEUU (www.data.gov).

De todas formas, no debemos caer en el error de pensar que este tipo de análisis será completamente automatizado. La intervención humana siempre será necesaria, ya que Big Data en el fondo pretende predecir el comportamiento humano, lo que requiere más la intervención de un antropólogo que de un analista de mercado. El comportamiento cultural humano se basa en aspectos como el deseo de mejorar el estatus, necesidad de relacionarse, ser competitivo dentro de un grupo, mientras que los estudios de mercado no tienen como objetivo estudiar estas características. 