

Big Data no tiene que ser necesariamente 'mejores datos'

Oscar Romero, director de ventas para **México y el Norte de América Latina para Infor** ha realizado un escrito de opinión que recuerda que la era digital es realmente la era de estar informados, “pero debido al fácil acceso a las plataformas sociales, **donde coexisten hechos reales, así como hechos falsos, tenemos acceso a mas información de la que es necesaria y útil**” y, con esto, el experto quiere analizar hasta qué punto los famosos ‘Big Data’ son también buenos datos.



Infor sostiene que es mejor que sobren datos a que falten. “Podemos decir que los datos especiales

mente big data, son similares al fuego, en la medida justa, utilizados de la manera correcta iluminan nuestro camino, disipan nuestras dudas y nos ofrecen calidez”. Pero, por contra, “si se utilizan de una manera incorrecta pueden ser destructivos y devastadores”.

Si consideramos la información que se usa para servir a los huéspedes, existen 4 tipos distintos de datos, dependiendo de su origen:

–**Activos**. Cuando específicamente solicitamos información de un huésped o cliente, eso es recolección activa de datos. Podemos **preguntar personalmente o a través de un formulario, pero de cualquier forma estamos pidiendo directamente lo que necesitamos para utilizar esos datos** en beneficio del cliente, o al menos esa es la teoría. Los datos activos son en general los mejores datos que se obtienen; (a menos que el que responde esconda la información o simplemente no desea brindar los datos solicitados), son datos precisos de la fuente original. Sin embargo, es importante destacar que hasta los datos activos tienen una vida útil: apellido, dirección, edad, estado civil, fumador o necesidades específicas, que pueden variar.

–**Pasivos**. Durante el curso rutinario de un hotel, se genera una enorme cantidad de datos. Contar con estos datos disponibles o utilizarlos para una toma de decisiones futura es un ejemplo de recolección pasiva de datos. Se puede asumir que al ser datos de hecho (durmió en una habitación

king, en el hotel xx), reflejarían las preferencias del huésped por una cama de ese tamaño. También es razonable pensar que se recolectó información indirectamente en el check in (prefiere cama doble o dos camas. Por lo que el cuarto doble será la probable referencia del huésped. Esto es así si realmente hubo una elección de parte del huésped y no que era lo único disponible en ese momento en el hotel.

-**Presuntos.** Los datos presuntos no son datos basados en recolección activa ni pasiva, sino porque **condicen con el perfil de una persona específica o grupo de afinidad. Los datos presuntos, se ajustan al dicho “Dios los cría y ellos se juntan”.** El riesgo de utilizar datos presuntos es que no son siempre reales, por ejemplo, se puede considerar que una persona por su lugar de origen tiene ciertas tendencias políticas o culturales. En un hotel según sus estadísticas si alguien se aloja un lunes o martes, consideran que pedirá cama doble. Sin embargo, existen los casos de personas que siempre piden camas separadas, una para dormir y la otra para apoyar la ropa, papeles etcétera. Los datos presuntos funcionan siempre que el cliente no sea justo la excepción.

-**Acumulado.** Otra forma de conseguir datos es recolectar y junta una gran cantidad de datos crudos, asumiendo que cuantos más datos se cuentan, más probabilidad de extraer información interesante. Este es el ingrediente secreto del big data, el total refleja más que las partes. Con estos datos se pueden hacer promedios, ver tendencias y evaluar procesos estadísticos.