

El 'big data' que todo lo sabe

El procesamiento masivo de datos sirve para el marketing personalizado en internet, pero también para estudios de genética, meteorología o la lucha contra el fraude

INVESTIGACIÓN

NOELIA A. ERAUSQUIN



«Vosotros no sois los clientes de Twitter o de Facebook, sois el producto». Así intenta explicar a sus alumnos el profesor del departamento de Informática de la Universidad de Oviedo, Daniel Gayo, que aquello que buscan o revelan en internet deja un rastro. No se trata de que alguien suba una foto más o menos comprometida y que esta circule por la red, sino que cada búsqueda o comentario va creando un perfil que luego es suministrado a terceros. ¿Y quiénes son esos terceros? Empresas que pueden beneficiarse de esa información y que son los clientes reales que hacen ganar cantidades ingentes de dinero a las compañías que están detrás de las redes sociales o buscadores. Esos datos les sirven, por ejemplo, para ofrecer publicidad adaptada a los gustos de cada persona o afinar las búsquedas. Así no es extraño que después de realizar consultas sobre viajes lleguen ofertas relacionadas con esos destinos. Se trata del 'big data', la capacidad de almacenar y procesar datos masivos.

Navegar por internet no es una actividad solitaria y hay muchos intermediarios interesados en conocer qué tipos de usuarios visitan qué sitios. Eso sí, en opinión de Gayo, especialista asturiano en la materia, no hay un Gran Hermano observando las acciones 'online' de cada uno, pero se va dejando un rastro disperso a través de distintos proveedores, una información que puede ser utilizada para ofrecer un servicio más ajustado a los intereses del internauta, pero también para otros asuntos más controvertidos como localizar disidentes o buscar posibles terroristas. «Estas tecnologías no son buenas o malas per se, sino en función del uso que se haga de ellas», aclara.

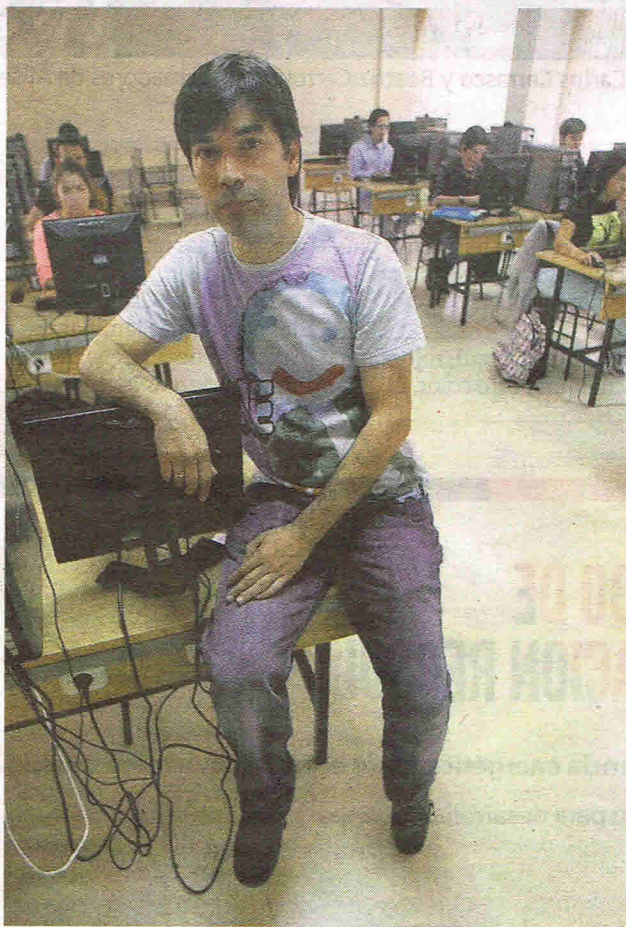
Cada vez que se sube un contenido a la red o se envía

en un correo empieza a ser explotado por la compañía que ofrece el servicio. No se trata de que alguien esté leyendo los comentarios o viendo las fotos, pero sí se analiza de manera automática. «Como diría el Padrino, no es personal, solo negocios. El problema es que para los usuarios la privacidad debería ser algo muy personal», destaca.

Lo cierto es que la mayoría de los usuarios no se preocupa de esta supuesta intromisión en su intimidad, aunque hay formas de luchar contra ella: divulgando menos información en las redes, desactivando el uso de 'cookies', utilizando buscadores sin estar validados como usuarios o incluso empleando mecanismos para no utilizar la propia dirección IP. Sin embargo, pocas personas toman estas precauciones, bien por desconocimiento o porque exigen un esfuerzo. Por otro lado, el catedrático de Informática Darío Álvarez recuerda que los ciudadanos están amparados por la Ley de Protección de Datos, independientemente del sistema técnico que se uti-



Ana Colubi y Gil González muestran cómo se puede representar los grandes bloques de datos en un ordenador. :: A. PIÑA



Daniel Gayo, en un aula de Informática. :: M. ROJAS

La mayoría de usuarios de internet no se preocupa por tomar precauciones para mantener su privacidad

lice, aunque de momento la normativa solo se aplica a empresas europeas. Pese a esto confía en que en el futuro la Ley cambie y sirva para proteger a los ciudadanos sin tener en cuenta la ubicación de la compañía.

Ana Colubi y Gil González, catedrática y profesor de Estadística de la Universidad de Oviedo, respectivamente, también trabajan con 'big data', aunque de otra forma. Perciben su fenómeno en internet con preocupación, «nos recuerdan que también es muy útil en otros ámbitos menos controvertidos, como el médico o la lucha contra el fraude. «Para un estadístico esto siempre se llamó 'large data

sets', pero le han dado otro nombre y se ha puesto de moda», señalan. «Ahora se genera más información de la que podemos procesar, las técnicas avanzan más lentas que la capacidad de almacenamiento», explica Colubi, que lidera el grupo internacional CMStatistics con más de 700 miembros y en el que varios grupos trabajan sobre esta temática.

Lejos de la percepción negativa que pueda tener para los usuarios de internet, estas técnicas son frecuentemente utilizadas para encontrar patrones genéticos de una enfermedad, para descubrir riesgos financieros o para prevenir la posibilidad de incendios o inundaciones, algo que sucede en Asturias. Otro caso es la lucha contra la evasión de impuestos. De hecho, la Universidad de Oviedo trabaja con otras entidades académicas europeas en una base de datos sobre las operaciones comerciales entre personas de países de la Unión. Con ella, buscando los comportamientos anómalos se podrán detectar casos de fraude.