

Nuevas fuentes y retos para el estudio de la movilidad urbana

JOAQUÍN OSORIO ARJONA¹ ✉ | JUAN CARLOS GARCÍA PALOMARES² ✉

Recibido: 08/12/2016 | Aceptado: 22/04/2017

Resumen

El aumento de la demanda de movilidad en las ciudades ha conllevado una dinámica poco sostenible tanto a nivel social como ambiental. Para promover actuaciones hacia una movilidad sostenible es necesario el uso de fuentes de información dinámicas, con un alto detalle espacial y temporal (y de bajo coste) que permitan realizar diagnósticos eficientes de la situación de movilidad en nuestras ciudades. Las Tecnologías de Información y Comunicación y el Big Data aparecen como nuevas fuentes interactivas que responden a estas necesidades. En este artículo se realiza una revisión del estado del arte en el uso de estas nuevas fuentes de datos para el análisis de la movilidad urbana, contrastando su utilidad respecto a las fuentes tradicionales, clasificándolas, presentando las temáticas de investigación que ofrecen, y abordando los desafíos de cara al futuro.

Palabras clave: movilidad, *Smart Cities*, TIC, redes sociales, *Big Data*.

Abstract

New sources and challenges for urban mobility studies

Increment of mobility demand in cities has carried socially and environmentally an unsustainable dynamic. For a sustainable planning, dynamic, high spatial and temporal resolution, and low cost sources are needed. Information and Communication Technologies appear as new interactive sources able to meet these needs. In this article, a state of the art in these new data sources use for urban mobility analysis is implemented. These new technologies are contrasted with traditional sources, are classified, new investigation topics are presented, and future challenges are addressed.

Keywords: mobility, *Smart Cities*, ICT, social networks, *Big Data*.

Abstrait

Nouvelles sources et challenges por l'étude de la mobilité urbaine

La demande croissante de mobilité dans les villes a conduit à une dynamique insoutenable tant au niveau social qu'environnemental. Dans le but de nous engager dans une approche durable, il s'avère indispensable d'utiliser des sources dynamiques d'information avec une haute précision spatiale et temporelle, et à faible coût. Les Technologies de l'information et de la communication émergent comme de nouvelles sources interactives répondant à ces besoins. Cet article examine l'état de l'art autour de l'utilisation de ces nouvelles sources de données pour l'analyse de la mo-

1. Departamento de Geografía Humana Universidad Complutense de Madrid. Grupo t-GIS. joaquoso@ucm.es

2. Departamento de Geografía Humana Universidad Complutense de Madrid. Grupo t-GIS. jgarcia@ucm.es

bilidad urbana. Ces nouvelles sources sont classifiées et leur utilité est contrastée par rapport aux sources traditionnelles. Enfin, de nouveaux sujets de recherche sont présentés et les enjeux d'avenir relevés.

Mots-clés: movilidad, *Smart Cities*, ICT, réseaux sociaux, *Big Data*.

1. Introducción

La movilidad es uno de los grandes retos de las ciudades del siglo XXI. Es entendida como la suma de los desplazamientos individuales de los ciudadanos, que posibilita su acceso al mercado de trabajo, bienes y servicios (Gasparini y Guidicini, 1990). El aumento del número de viajes, la mayor diversidad de motivos de desplazamiento, el uso intensivo de transportes motorizados, y los recorridos cada vez más largos han creado una situación poco sostenible, tanto a nivel social como ambiental. Nos encontramos ante un fenómeno multivariado y de diversificación de horarios como consecuencia de diferentes modelos individuales de movilidad (Banister, 2011). Ante esta situación, es necesario plantear un modelo de gestión sostenible, que permita un consumo eficiente y equilibrado del transporte, no consuma materia y energía de manera excesiva y facilite la reducción de la contaminación mediante un mayor uso de desplazamientos peatonales o en bicicleta, y una planificación más inteligente del transporte público (Banister, 2008).

La necesidad de un modelo sostenible requiere de fuentes de datos que permitan un conocimiento profundo de la movilidad urbana. Es clave estudiar los diferentes comportamientos de movilidad de los ciudadanos, adquiriendo las fuentes un papel crucial para ello (Schwanen, 2016). La información sobre pautas de movilidad tiene una importancia vital, pues de ella depende la elaboración de diagnósticos correctos, y la utilización de las herramientas de modelización y la estimación de los impactos de las actuaciones a promover (Miralles-Guasch, 2012; Miralles-Guasch y Martínez, 2013). Sin embargo, las fuentes que proporcionan información sobre movilidad son tradicionalmente escasas. Habitualmente los datos utilizados por los gestores del transporte provienen de Encuestas Domiciliarias de Movilidad (EDM). Estos datos son de gran interés, pues proporcionan información detallada y específica de la movilidad, pero a la vez tienen costes elevados, lo que reduce la temporalidad (frecuentemente diez años) o el tamaño de las muestras utilizadas, conllevando niveles de detalle espacial menor. Se hace necesario disponer de otras fuentes de datos, que proporcionen información complementaria y con una escala espacial y temporal mucho mayor. Todo ello, a ser posible, con costes económicos menores.

El objetivo de este artículo consiste en realizar una revisión del estado del arte sobre nuevas fuentes de datos para el estudio de la movilidad urbana, centrándose en el papel de las Tecnologías de la Información y la Comunicación (TIC) en el contexto de las *Smart Cities*, como herramientas interactivas y dinámicas que permiten obtener información con gran frecuencia, habitualmente en tiempo real. Se atenderá a las posibilidades que abren las fuentes asociadas al llamado *Spatial Big Data* (SBD): bases de datos masivos, que recopilan información con atributos espaciales, temporales y otros campos que ayudan al desarrollo de la investigación en el campo de la movilidad urbana. Al estar condicionada a las nuevas tecnologías que van surgiendo y evolucionando en los últimos años, esta temática ha sido implementada muy recientemente. A nivel internacional hay pocos trabajos que traten este tipo de fuentes y son escasos los trabajos en el ámbito nacional.

Este artículo comienza efectuando una breve recopilación de las fuentes de datos que se han usado habitualmente para obtener información relacionada con la movilidad y se comparan con las

procedentes de las TIC y el *Big Data*. A su vez, se estudiarán las características y ventajas de las nuevas fuentes frente a las bases de datos tradicionales para el estudio de la movilidad. Después, se entrará en detalle en los distintos tipos de nuevas fuentes de datos, destacando los dispositivos móviles, las redes sociales y las *smart cards* de transporte público. Se contemplarán las oportunidades que el SBD tiene para la investigación, junto con nuevas temáticas y aplicaciones que ha abierto su uso. Finalmente se presentan retos que ofrecen a la gestión y tratamiento de la información para el estudio de la movilidad.

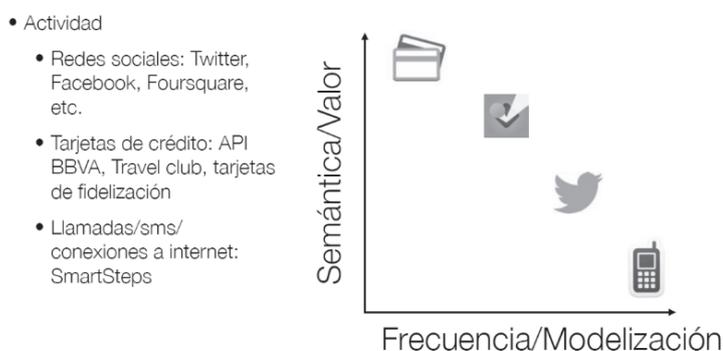
2. Metodología

La metodología empleada para la revisión del estado del arte ha sido la búsqueda bibliográfica en las bases de datos de *Google Scholar* y *Scopus*, donde se ha tratado de buscar todos aquellos trabajos en los que se estudia la movilidad a partir de las nuevas fuentes de datos. Inicialmente se ha realizado también una revisión de fuentes de datos tradicionales. En este caso, dada la cantidad de bibliografía existente, la búsqueda se ha centrado en trabajos de revisiones de la literatura sobre fuentes de movilidad tradicionales.

En la consideración de las nuevas fuentes de información, la atención se ha puesto en aquellas que tienen que ver con el uso de la telefonía móvil, las tarjetas inteligentes de transporte y tarjetas de crédito, y en especial las vinculadas a las redes sociales en internet. Estas últimas están cobrando una importancia muy alta en el estudio de la movilidad, donde se están usando cada vez más. Se ha empleado la clasificación que Esteban Moro (2016) realizó de los nuevos proveedores de datos (Imagen 1), en función de la frecuencia con la que se produce información y el valor de la información que puede obtenerse de las mismas. Así, la geolocalización de llamadas de teléfono o de mensajes de texto en Twitter (*tweets*), producen información con una frecuencia muy alta pero la información que puede obtenerse de las mismas es de menor valor. De estas fuentes habitualmente apenas se puede obtener las propias coordenadas, el momento temporal o alguna característica socioeconómica asociada a los usuarios. Por el contrario, las tarjetas de crédito se usan con una frecuencia menor, pero el valor de la información que proporcionan es mucho mayor.

Imagen 1. Fuentes de información por frecuencia y valor

Proveedores de datos



Fuente: Moro (2016)

Una vez definidas las fuentes de datos sobre las que se va a realizar la revisión bibliográfica, se presentan las líneas de trabajo que tienen que ver con el uso de estas fuentes. La metodología utilizada para realizar esta revisión ha sido la búsqueda, lectura y análisis de situaciones prácticas de empleo de las nuevas fuentes de datos en ámbitos de movilidad urbana. Para ello se han revisado los trabajos aparecidos en revistas científicas y también, dada la novedad y actualidad de la temática, en conferencias y seminarios de investigación. Se han señalado especialmente ejemplos que permitan comparar y evaluar las nuevas herramientas de fuentes de datos sobre los métodos más tradicionales. A su vez, se han seleccionado investigaciones que hayan aportado nuevos métodos y aplicaciones metodológicas de interés.

Revisados estos trabajos, se ha optado por agruparlos y ordenarlos según su objetivo planeado: definición de espacios de atracción y generación de viajes, diseño de matrices origen-destino de viajes, identificación de pautas de comportamiento individual, observación de impactos de eventos en la movilidad urbana, y uso de la información para conocer la percepción de los servicios de transporte.

3. Fuentes tradicionales y nuevas fuentes de datos

3.1. Fuentes tradicionales

A nivel local, suele producirse un vacío legal sobre las competencias en cuanto a recopilación, tratamiento y difusión de la información sobre movilidad, por lo que las fuentes de información han dependido generalmente del departamento de tráfico o transporte correspondiente. En las principales ciudades, existen datos de flujos de tráfico urbano. Sin embargo, los datos oficiales disponibles para el investigador suelen ser escasos.

Tradicionalmente las Encuestas Domiciliarias de Movilidad (EDM) se han usado para recopilar información del origen y destino de los viajes, los medios de transporte empleados y sus etapas, duración, hábitos de transporte de la población y otros datos relevantes (Griffiths y Richardson, 2000; Miralles-Guasch, 2012; Miralles-Guasch et al., 2015; Ortúzar S. y Willumsen, 2011). Además, recogen gran cantidad de información sobre la población, permitiendo conocer con detalle sus características sociodemográficas y económicas, así como información relevante sobre la vivienda. Al tratarse de una fuente específica para el estudio de la movilidad y para alimentar el desarrollo de modelos de transporte, la riqueza de la información obtenida es amplia. Sin embargo, las EDM también presentan inconvenientes: su modelización requiere de enormes muestras de hogares encuestados y adaptados a zonificaciones de transporte, un considerable trabajo de campo, y un gran esfuerzo personal, lo que eleva los costes para su realización. Estos costes acaban reduciéndola periodicidad de las encuestas y la resolución espacial, que implica en ocasiones zonas de transporte de gran tamaño. Otras limitaciones consisten en viajes no reportados, tiempo impreciso y falta de información sobre localizaciones o rutas (Zhao et al., 2015).

Otro tipo de fuente habitual son las encuestas panel sobre movilidad, que recogen información de una muestra de población durante un periodo determinado de tiempo (por ejemplo, una semana), y en distintos momentos temporales. De esta forma se persigue recopilar información sobre cambios en patrones de movilidad, normalmente como consecuencia de algún tipo de actuación en materia de transporte. Nuevamente, la larga duración de este tipo de encuestas, que obliga a mediciones sucesivas e intervalos regulares, eleva su coste.

Por otro lado, también se realizan conteos y aforos en los distintos medios de transporte, como estimaciones estadísticas representativas de su uso. Para ello se obtiene una muestra aleatoria dentro de un ámbito concreto, que es computada estadísticamente, resultando en estimaciones rápidas y de bajo coste. Sin embargo, difícilmente se obtiene información sobre la población y sus características.

En otros casos se utilizan técnicas cualitativas, como diarios de viaje, observaciones, entrevistas, grupos de discusión, etc. Estas técnicas son métodos más accesibles sobre la movilidad, basándose en la interacción directa con la población objeto de estudio (Pazos, 2005). Se han usado con cierta frecuencia para analizar la movilidad de determinados grupos de población.

También se puede recoger información de fuentes indirectas, como los propios censos de población, para obtener datos a partir de cuestionarios sobre los lugares de trabajo y vivienda y del medio de desplazamiento utilizado (García-Palomares, 2010). Sin embargo, estos censos cuentan cada vez con menor información, su periodicidad es baja y su coste nuevamente es elevado.

Miralles-Guasch y Martínez (2013) han evaluado la valoración que realizan los profesionales en la planificación y gestión del transporte de las fuentes de datos tradicionales de movilidad urbana. Se muestra como, a pesar del interés de estas fuentes, actualmente son métodos de obtención de información poco dinámicos, que requieren de una constante actualización para tener información adecuada, haciendo que sean costosas y difíciles de mantener. Además, conllevan un tiempo muy importante de preparación (por ejemplo, con la preparación de las zonificaciones de transportes o las estimaciones de las muestras), afectando directamente al coste y datos obtenidos. También, la baja resolución tanto temporal y espacial hacen que el tipo y cantidad de datos obtenidos sea limitado, dando en ocasiones como resultado bases de datos escuetas e incompletas (GHD, 2014). Por último, las bases de datos de estas fuentes sufren una progresiva reducción de la cantidad y la calidad de sus contenidos (Miralles-Guasch et al., 2015). El avance tecnológico abre oportunidades para que este tipo de fuentes puedan ser complementadas con información procedente de otras fuentes, vinculadas al desarrollo de las TIC y las redes sociales, las *Smart Cities*, o el *Spatial Big Data*.

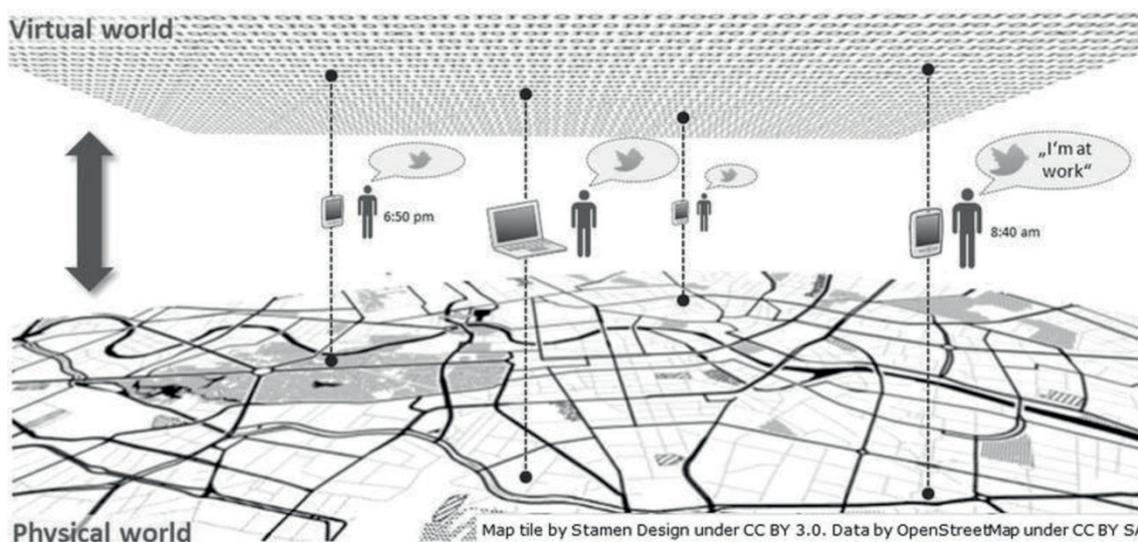
3.2. Nuevas fuentes de datos

Las nuevas fuentes de datos aparecen como medios potentes y dinámicos, fácil es de tratar y que ofrecen la posibilidad de hacer análisis más profundos y complejos, donde se complementa la información obtenida de encuestas de movilidad. Las TIC engloban un conjunto de herramientas (redes, terminales, servicios) que pueden ser activas o pasivas dependiendo de si el usuario es consciente o no de la propia generación de información (Miralles-Guasch et al., 2015). De muchas formas distintas, las personas y las cosas, sobre todo en las ciudades, están monitorizadas continuamente, generando datos, muchas veces georreferenciados y que se almacenan en diversos repositorios, públicos y privados, por lo que se produce una continua explosión de datos (Bosque, 2015; Gutiérrez et al., 2016).

Recientemente los expertos están asegurando que el binomio formado por Internet y el *Big Data* están configurando una cuarta revolución industrial, la denominada revolución de la información. Para entenderlo, hay que verlas TIC como el nodo de enlace entre el mundo virtual, donde se halla la información, y el mundo físico, donde se proyecta dicha información espacial y temporalmente (Imagen 2). Se puede decir que los datos son impulsos neuronales de un sistema ner-

vioso digital que recubre poco a poco el mundo físico con sentidos y músculos, evolucionando las ciudades a un entorno cada vez más inteligente (Tascón y Coullaut, 2016). Asociada a esta revolución surge el concepto de las *Smart Cities* cómo ciudades que se adaptan para alcanzar un desarrollo eficiente y sostenible, una participación ciudadana activa, y una mejor calidad de vida. El *Big Data* enriquece el conocimiento acerca de las ciudades y ofrece nuevas oportunidades para la interacción social y la planificación (Batty, 2013).

Imagen 2. Las TIC como enlace entre el mundo físico y el virtual



Fuente: (Steiger et al., 2015)

El *Big Data* y las TIC se antojan como nuevas fuentes al servicio de la ciudad. Con la *Web 3.0* ha surgido el concepto de *Spatial Big Data* (SBD): grandes cantidades de información espacial que pueden ser capturadas, comunicadas, agregadas, almacenadas y analizadas (McKinsey & Company, 2011). Estas fuentes pueden recolectar datos constantemente, frecuentemente en tiempo real, siendo herramientas dinámicas que permiten la actualización progresiva de la información. Además, la facilidad de disponibilidad y de obtención del equipamiento necesario hace que el coste sea normalmente bajo, pudiendo obtenerse y actualizarse en muchos casos información de forma gratuita.

El *Big Data* posee tres características principales: volumen enorme, velocidad alta y variedad amplia. Dichas características son esenciales a la hora de manejar fuentes relacionadas con la movilidad. El volumen está ligado al crecimiento del tráfico de datos y de sus fuentes de información. La velocidad se vincula al ritmo acelerado al que se producen los datos y la velocidad de las tecnologías y de sus funciones de monitorización y procesado. La variedad se debe al amplio espectro de fuentes desde las que se producen (Clarke, 2016).

En relación con la movilidad, se puede clasificar la información respecto a si estudia las infraestructuras, los vehículos o los usuarios. Con el *Spatial Big Data* se pueden ver los atributos de las infraestructuras, si están operativas, si necesitan mantenimiento, se pueden formular previsiones, o diseñar mapas dinámicos de su rendimiento. Con los vehículos, es posible monitorizar la localización, el rendimiento, el conductor, o realizar telemetrías. En cuanto a los usuarios se contemplan su estado, actividad personal, y comportamiento (Europe, 2015).

Las TIC tienen además una alta interoperabilidad con las fuentes tradicionales, permitiendo muchas veces crear encuestas. Esta complementariedad otorga a las fuentes tradicionales un alcance del que no suelen contar, y una mayor flexibilidad. Además, las TIC permiten obtener bases de datos de alta resolución espacial, con datos geolocalizados a nivel de punto que permiten el tratamiento de información muy completa y la realización de análisis a diferentes escalas, desde el nivel global hasta el de la localización concreta. Por último, no están limitadas a capturar información de un espacio concreto, sino que puede conseguir datos con una cobertura global, lo que permite realizar estudios comparados entre ciudades.

El Spatial *Big Data* es fácil de implementar en Sistemas de Información Geográfica (SIG), permitiendo realizar análisis espaciales y visualizaciones dinámicas de resultados. Debido a la abundancia y amplitud de información se puede diseñar cartografía de diferentes temas (Ciuccarelli et al., 2014). Además, los propios SIG se han convertido en una herramienta social aprovechada por las TIC para contar historias y noticias (Sui y Goodchild, 2011). En el campo teórico de la Geografía, cuenta con un gran potencial en forma de información abundante y por la posibilidad de promover su valor académico a una mayor audiencia (Kitchin, 2013).

Gracias a las nuevas fuentes y sus análisis en entornos como los SIG, es posible ampliar las miras en estudios de movilidad desde una doble perspectiva. Por un lado, permiten obtener mejores respuestas a problemas ya planteados, con mayor detalle a partir de las posibilidades que ofrece la mayor resolución espacial y temporal de los datos. Por ejemplo, frente a los modelos para estimar pautas de movilidad o de transporte futuros (*forecasting*), han empezado a surgir herramientas apoyadas en *Big Data* para realizar esas predicciones en tiempo real (*nowcasting*) (Hanabusa, 2012). Por otro lado, debido a la posibilidad de adquirir información con atributos sobre diferentes aspectos que no era posible capturar con los métodos tradicionales, es posible ampliar las temáticas de estudio. Surge la oportunidad de estudiar temas que no habían sido tratados o que habían sido dejados de lado ante la falta de información. Es el caso de temas como la movilidad turística o la gestión de eventos.

4. Nuevos tipos de fuentes para el estudio de la movilidad urbana

La constante evolución, mejora y actualización de las TIC han contribuido a generar nuevas fuentes para el estudio de la movilidad. El continuo desarrollo de Internet y de plataformas que permiten su uso móvil permite crear y almacenar información compartida de forma instantánea, dando lugar a un importante flujo de datos que se distribuye y se hace accesible a cualquier lector desde cualquier punto del planeta (Gutiérrez-Puebla et al., 2016; Kitchin, 2013). En este apartado se tratarán las nuevas fuentes según la clasificación de Moro (2016).

4.1. Teléfonos móviles, Wireless, y navegadores

El acceso a los teléfonos móviles o *smartphones* es cada vez mayor y más sencillo, viviéndose una auténtica proliferación en la que los teléfonos móviles son ya parte indispensable de la vida diaria de la mayor parte de la población metropolitana. Un ejemplo de su función en el comportamiento de movilidad urbana radica en la incorporación de aplicaciones diseñadas por empresas de transporte para informar al usuario de los tiempos de llegada o paradas disponibles en el recorrido a realizar. Con el aumento de su funcionalidad y complejidad, los *Smartphones* cuentan con aplicaciones que sirven para infinitas funciones. Entre ellas está la posibilidad de recoger la ubica-

ción del usuario gracias a la función de sistema de posicionamiento global (GPS) o la localización de los puntos de la red telefónica a los que se está conectando. Con este sistema se puede indicar la localización de la persona o la ruta de un determinado vehículo. A partir de la geolocalización de las llamadas, podemos identificar patrones de movilidad de diferentes grupos de población, las infraestructuras más utilizadas, o los espacios más transitados (Lathia et al., 2013). El uso de teléfonos móviles para determinar los patrones de movilidad de un individuo conlleva un acercamiento mucho más preciso gracias a la capacidad de monitorear la actividad física de una persona, o los perfiles de movilidad de un determinado modo de transporte (Reddy et al., 2010).

Otras herramientas son los sensores que detectan vehículos o usuarios a través de tecnología *wireless*. Entre ellos destacan sensores acústicos, radares de microondas, sensores de infrarrojos o *LIDAR*, dispositivos *Bluetooth*, o etiquetas de identificación de frecuencia por radio (GHD, 2014). A su vez, los sistemas de detección de imagen por video se pueden aplicar para realizar conteos de peatones mediante cámaras especializadas, siendo una aplicación muy útil para el estudio de aglomeraciones en determinados lugares.

Finalmente, uno de los tipos de dispositivos que produce información de gran interés para el estudio de la movilidad es el compuesto por los sistemas de navegación con los que están equipados hoy en día la mayoría de los propios vehículos. Estos sistemas indican al conductor en un mapa su ubicación, los servicios más cercanos o el recorrido óptimo a efectuar en su trayecto, recogiendo a la vez dicha información. Gracias a la base de datos del navegador que almacena las rutas efectuadas es posible hacer una recopilación de trayectos realizados por una persona a lo largo de un tiempo determinado y obtener otros datos de interés como el tiempo empleado. En los últimos años han surgido compañías que utilizan este tipo de herramientas como *Google Transit*, *Uber*, *BlaBlaCar*, u otras muchas similares.

4.2. Web 3.0: redes sociales

El concepto de *Web 2.0* constituyó un avance importante en el uso de Internet, al conferir a la red la facilidad de compartir información y la interacción. Desde hace pocos años se habla ya de la *Web 3.0*, cuyos pilares son la interoperabilidad, las tecnologías semánticas y el ambiente de computarización social. La idea básica se halla en definir la estructura de datos y enlazarlos de modo que su uso sea más eficaz y accesible, siendo las bases de datos enlazadas una de las claves en las que radica también el *Big Data* (Sareh et al., 2012). Una de las herramientas fundamentales de Internet y las TIC son las redes sociales, que permiten crear, almacenar, compartir e intercambiar información con otros usuarios (Cao et al., 2014). Su función general consiste en la interacción con otras personas a través de diferentes medios, generándose diariamente una importante cantidad de datos. Muchas empresas usan las redes sociales como recursos para obtener información, estudiar campos de actuación, o publicitarse. En muchos casos la información de las redes sociales está geolocalizada, y tiene un carácter dinámico debido a su constante actualización, lo que permite conocer las pautas de distribución de los usuarios y su movilidad. En el caso del estudio de la movilidad, las redes sociales son valiosas para informar acerca del tráfico a distintas horas del día, compartir horarios o eventos, dar a conocer empresas y modos de transporte, o planificar proyectos que afecten a distintas áreas urbanas.

Una de las redes sociales más populares es *Twitter*, por su sencillez y facilidad para usarse desde cualquier dispositivo. Esta aplicación está basada en el *microblogging* que permite a sus usuarios enviar pequeños textos ("*tweets*"), de un máximo de 140 caracteres, además de otros contenidos.

Los *tweets* pueden ser geolocalizados a partir de coordenadas GPS de los *smartphones*, recogiendo en el mensaje la localización desde donde se realiza. *Twitter* es una red social inmensamente popular, que genera más de 500 billones de *tweets* al día. Moro (2014)³ reproduce la movilidad de los usuarios de *Twitter* en España durante un día, realizada a partir de 145 millones de mensajes entre 2012 y 2013, lo que permite apreciar la intensidad del uso de esta red social.

Twitter se ha vuelto una herramienta ideal y original para estudiar la movilidad y los patrones de comportamiento de los flujos de un área metropolitana. Gracias a la localización del *tweet*, es posible conocer la distribución de la población en cualquier momento temporal, pudiendo establecer comparaciones entre diferentes fechas y a la vez comprender la evolución que sufre un área metropolitana a lo largo de un tiempo determinado (día, mes, año, etc.). A la vez, el acceso instantáneo a opiniones, críticas y datos a tiempo real, localizados tanto temporal como espacialmente, se antoja de un valor clave a diferentes sectores que usan *Twitter* como herramienta para conducir encuestas y estudiar la opinión de usuarios o grupos sobre el funcionamiento de los modos de transporte (Luong y Houston, 2015).

Otras redes sociales de gran interés son las basadas en la geolocalización de fotografías, siendo fuentes fáciles de combinar con otras redes sociales. Las comunidades más importantes son *Instagram*, *Flickr* y *Panoramio*, que permiten visualizar fotografías etiquetadas espacialmente en mapas. Estas fuentes permiten conocer los espacios más fotografiados, y en consecuencia los más visitados, siendo posible también recoger imágenes de un trayecto y mostrarlas cartográficamente a escala local (García-Palomares et al., 2015). También destaca *Foursquare*, red social que permite a través del GPS compartir con otros usuarios lugares específicos marcados en mapas por el propio individuo, lo que le convierte en un programa ideal para el estudio de puntos de reunión, zonas con mayor tránsito o atracción turística, o actividades basadas en el tipo de equipamiento compartido.

4.3. Tarjetas inteligentes de transporte público y tarjetas bancarias

Un tipo de fuente de información de gran interés en el estudio de la movilidad son las tarjetas inteligentes o *smartcards*. Estas tarjetas cuentan con un circuito integrado microprocesador que permiten tanto la identificación del titular como el almacenamiento de información asociada. Las tarjetas inteligentes de transporte son bastante usadas en transportes públicos (autobús, metro o tren), siendo muy valiosas para el estudio de la movilidad debido a su fácil adquisición, su alta demanda, y su constante uso. Al ser pasadas por un programa lector, permiten el almacenamiento de patrones de viaje de sus usuarios, siendo un modo muy útil de analizar dinámicas urbanas a corto, medio y largo plazo. Entre la información que se genera destacan el medio de transporte utilizado, la parada de origen, y en algunos casos la parada de destino, los trasbordos efectuados, y el tiempo empleado en el viaje (Cao et al., 2014).

Las tarjetas bancarias también permiten ubicar a su portador en los dispositivos y cajeros donde se hayan usado, incluyendo máquinas facturadoras de billetes de transporte público que aceptan pago por dicho medio. La georreferenciación de cada una de las transacciones bancarias realizadas por los usuarios de las tarjetas permite realizar un seguimiento de los individuos y conocer su movilidad en la ciudad (Lenormand et al., 2015).

3. <https://vimeo.com/111579945>

5. Oportunidades para la investigación de la movilidad urbana: temáticas y aplicaciones

En los últimos años aparece un uso incipiente del *Spatial Big Data* para crear bases de datos con las que estudiar diversas problemáticas relacionadas con la movilidad. A continuación, se citan distintas líneas de trabajo, que tienen que ver con el uso de estas fuentes.

5.1. Definición de espacios de atracción y generación de viajes

Las fuentes que tienen una gran frecuencia de uso permiten conocer el uso del espacio urbano y, definir espacios de atracción y generación de viajes. Por ejemplo, en Manhattan se crearon mapas de actividades a partir de las distribuciones horarias en el uso de *Twitter* y se compararon con el mapa oficial de usos del suelo de Nueva York (Frias-martinez et al., 2012). También se ha usado *Twitter* para observar el uso del suelo en Londres, diferenciando según grupos demográficos (grupos étnicos, por género o por edad). Para verificar los resultados obtenidos, éstos se comparan con los del censo de población (Longley y Adnan, 2016).

Estudios similares se realizaron para ver distribuciones demográficas espacio-temporales de la ciudad de Chicago, y construir trayectorias con las que poder identificar centros de actividad o residencia, complementadas con un filtrado de *tweets* por grupos con el cual poder realizar análisis diferenciados para cada uno de ellos (Luo et al., 2016). En Río de Janeiro se combinó la identificación de patrones espaciales con el censo para realizar una cartografía de distribución de población por poder adquisitivo (Netto et al., 2015). Mientras, en Louisville (Kentucky), se estudió los espacios donde realizan las actividades dos grupos étnicos diferentes, a partir de la distribución de *tweets* a diferentes escalas temporales (Shelton et al., 2015).

En la Escuela de Ingeniería Civil de la *Purdue University* se ha estudiado la movilidad y actividad urbana a partir de *tweets* geolocalizados en las ciudades de Nueva York, Chicago, y Los Ángeles, clasificando actividades a partir del tipo de localizaciones visitadas. Como resultado, observaron pautas de movilidad individual a nivel temporal analizando las distribuciones por actividades a diferentes horas del día (Hasan et al., 2013). Mientras, en Londres se han analizado cartográficamente las diferencias de densidades entre el día y la noche para estudiar distintos patrones de movilidad (Lansley y Longley, 2016).

En esta línea de trabajo, las *smartcards* también han sido utilizadas para observar el número de residentes con bajo poder adquisitivo que usa el transporte público en la ciudad china de Beijing e identificar trayectos de acuerdo con el nivel socioeconómico (Long y Shen, 2015). La metodología consistió en la adquisición de datos a lo largo de una semana de las tarjetas de transporte público, la identificación de los lugares de vivienda y de trabajo de sus usuarios, y la clasificación en grupos según las características del lugar de residencia o de trabajo. En una línea parecida, en Shanghai se buscó reproducir patrones de movilidad humana aprovechando los *check-in* de lugares de encuentro en páginas web para la categorización de demandas de viaje en diferentes categorías (Wu et al., 2014).

5.2. Aforos y matrices Origen-Destino de viajes

La gran frecuencia de información que proporcionan las nuevas fuentes de datos es muy útil para obtener aforos y mediciones de matrices de viajes. Un ejemplo es la estimación de aforos de trá-

fico a partir de teléfonos móviles en las vías de acceso a la ciudad de Sevilla (Caceres et al., 2012). Los resultados han sido comparados con aforos reales, mostrando que las mediciones realizadas tienen ajustes buenos en relación al aforo real. En Nueva Zelanda se calcularon aforos y mediciones del flujo peatonal, seleccionando puntos concretos en ciudades y colocando videocámaras. El número de peatones, sus características y la velocidad a la que caminan se calcularon a partir de la primera imagen recogida del peatón hasta el momento en el que sale de la visión de la cámara. Los resultados precisaron las velocidades de las personas observadas en función no solo de características básicas como el género, sino otras como el equipamiento o el tipo de calzado (Finnis K y Walton, 2007).

Un paso más es la obtención de matrices origen-destino (OD). La Universidad de California ha empleado *Twitter* para calcular matrices en el área metropolitana de Los Ángeles y en intervalos de 30 minutos. Para ello utilizaron un algoritmo basado en detección de trayectorias individuales y agregación de viajes a los diferentes destinos. Si una persona realizaba un *tweet* en diferentes zonas de análisis de tráfico y en un determinado margen de tiempo, se consideraba un viaje entre esas dos áreas. (Lee y Lee, 2015).

Otras fuentes de datos para la obtención de matrices OD son los sistemas de navegación de vehículos. En Southampton se realizó un estudio a través de una matriz obtenida con los datos originados por el transporte público, y subidos a la red por la propia empresa para facilitar la accesibilidad a equipamientos sanitarios en la región rural situada entre las ciudades de Devon y Cornwall (Martin et al., 2008). En Brisbane (Australia), se ha realizado una investigación similar a partir de *smartcards* para reconstruir trayectos de pasajeros de autobús en el periodo de un día. Con la construcción de estas matrices se pueden observar dinámicas de comportamiento espacio-temporal en el uso del transporte público (Tao et al., 2014). La Universidad de Gent ha usado datos de tiempos de transporte calculados en sistemas de navegación de coches para estudiar los comportamientos de movilidad que se dan en la región belga de Flandes. Con esta información se crearon dos matrices OD: una basada en el número de viajes en hora valle y otra para la hora punta (Dewulf et al., 2015).

Mientras, un equipo de Raleigh (Carolina del Norte) ha empleado navegadores GPS en teléfonos móviles para recopilar diferentes encuestas que contengan información de recorridos efectuados a diario en un periodo de 60 días. A continuación, crearon una base de datos para poder obtener matrices OD con información enriquecida que permitía clasificar diferentes rutas de movilidad según el transporte empleado o el momento del día (Badger, 2012).

La calidad de los nuevos datos para el estudio de matrices de movilidad urbana ha sido validada en varios estudios. En Madrid se compararon datos de matrices OD obtenidas a partir de datos de *Twitter* y de llamadas de teléfonos móviles con los datos oficiales del Censo. Los datos de correlación entre los resultados de las tres fuentes fueron significativamente elevados (Lenormand et al., 2014). También en Los Ángeles se compararon matrices obtenidas con información de censos oficiales para descubrir variaciones en los resultados obtenidos (Gao et al., 2014).

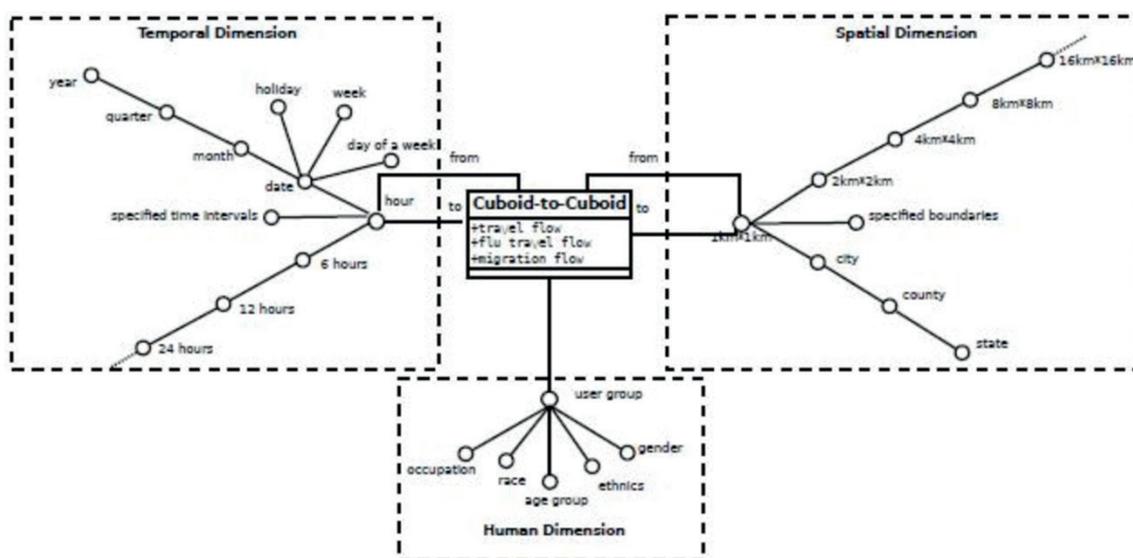
5.3. Pautas y comportamientos de movilidad individual

La alta resolución espacio-temporal de las nuevas fuentes de datos permite dar un salto al estudio de la movilidad individual. Por ejemplo, ha sido posible capturar el movimiento de la movilidad en Kenia a diferentes escalas espaciales y temporales gracias a que cada persona deja una huella

digital única con cada *tweet* que envían (Blanford et al., 2015). En Leeds, se han recogido *tweets* de la ciudad inglesa durante un año. Tras filtrar los mensajes de cuentas automáticas o de usuarios con pocos mensajes, realizaron un cálculo temporal a partir del tiempo transcurrido entre *tweets* empleados por una misma cuenta. A continuación, analizaron localizaciones de cada usuario y el tiempo total en el que han estado en cada una de ellas. En la siguiente parte del análisis estudiaron la proporción de tiempo gastado en casa, el nivel de actividad, y el balance de *tweets* entre días de trabajo y fines de semana. Todo ello permitía clasificar a los usuarios en diferentes categorías según sus características de movilidad (Birkin et al., 2014).

Para estudios con diversos grupos sociales, el concepto de información tridimensional es una herramienta eficaz, definiéndose la dimensión temporal cómo la duración y distribución horaria de las actividades individuales, la dimensión espacial cómo las configuraciones localizables de los patrones de actividad humana, y la dimensión humana cómo las interacciones sociales con otros individuos. Partiendo de este teorema se ha desarrollado el concepto de *datacube* para el análisis espacio-temporal de la localización social (Imagen 3). Esta herramienta consiste en una base de datos tridimensional donde se pueden examinar las dinámicas de un individuo en las redes sociales a múltiples niveles de escalas espacio-temporales y a diferentes reglas de agregación. Una vez construido el *datacube* se puede proceder a la construcción de trayectorias de los usuarios de *Twitter* y a la realización de análisis visuales y espacio-temporales en un SIG, permitiendo también crear un interfaz de cartografía *online* para representar las dinámicas de movimiento a partir de las trayectorias creadas (Cao et al., 2014).

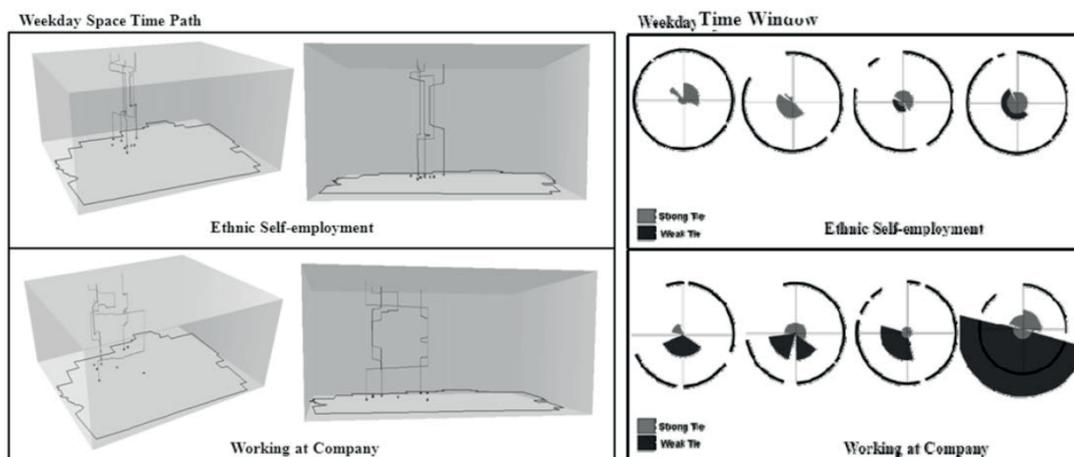
Imagen 3. Esquema de un *datacube*



Fuente: (Cao et al., 2014)

Utilizando metodología 3D, en la ciudad estadounidense de Columbiase usaron redes sociales para detectar fenómenos de aislamiento de grupos sociales particulares en el ámbito metropolitano, en este caso ciudadanos coreanos. Una vez recopilada la información, se visualizaron en 3D recorridos espaciotemporales, superficies de densidades creadas con estimaciones *kernel*, ventanas de tiempo basadas en relojes de 24 horas, todo ello reflejado en círculos concéntricos que mostraban diferentes actividades en anillos de una hora (Imagen 4) (Kwan y Lee, 2011).

Imagen 4: Representación 3D de rutas espaciotemporales y ventanas de tiempo



Fuente: (Kwan y Lee, 2011)

En otro trabajo, para conocer la movilidad individual en bicicleta en la ciudad de Madrid, se han recogido los recorridos a través de una aplicación instalada en *smartphones* de ciclistas voluntarios, creando un mapa denominado “*La huella ciclista de Madrid*”. Este mapa permite conocer las calles más transitadas, y en consecuencia las zonas que necesitan mayor inversión en infraestructuras ciclistas (Romanillos y Zaltz Austwick, 2015). En Londres se han combinado datos de sistemas de navegación de taxis, información de ocupación de estaciones de bicicletas públicas, y *smartcards* de transporte público para poder realizar clústeres que revelen comportamientos emergentes a nivel individual, y así poder comparar los resultados obtenidos con otros estudios o con los tiempos de tránsito publicados por las propias empresas (Lathia et al., 2013).

5.4. Impactos de eventos en la movilidad urbana

La celebración de determinados eventos en nuestras ciudades produce un gran impacto en la movilidad, que debe gestionarse con la mayor eficiencia. Para ello, el empleo de fuentes que proporcionen datos en tiempo real es fundamental. En Países Bajos se ha empleado *Twitter* para la detección de eventos (Steur, 2014). Se sabe que hay un tiempo de retardo entre que se produce un incidente y que los centros de mantenimiento son informados, por lo que se ha buscado usar la información que suben los usuarios a *Twitter* en tiempo real con el objetivo de tratar de reducir ese tiempo mediante la identificación de la localización y el tipo de incidente producidos. La metodología parte de la recolecta e identificación de los *tweets* más relevantes para hacer un análisis zonal que permita identificar irregularidades a través de clústeres de *tweets* relacionados con el tráfico.

El Grupo de Investigación *GIScience* de la Universidad de Heidelberg ha usado un método de combinación de análisis semántico y análisis espacial para extraer también flujos de movilidad con *Twitter*. Una vez procesada la información buscaron encontrar similitudes semánticas entre *tweets* y a continuación efectuar un análisis de autocorrelación espacial para hallar patrones espaciotemporales (Steiger et al., 2015). Esta metodología ha sido empleada para descubrir patrones de movilidad durante eventos deportivos de masas en Boston, a partir de clústeres espaciales a diferentes horas del día.

Las tarjetas de transporte también se han utilizado para la predicción de comportamientos del uso del transporte público bajo eventos especiales en la ciudad de Singapur. En este caso, se seleccionó una serie de lugares de encuentro y se comprobaron en los datos conseguidos por las *smartcards* el número de llegadas que ocurrían para cada uno de esos lugares, con el fin de obtener un modelo estadístico de predicción aplicable a futuros eventos (Pereira et al., 2015). En Gent se ha realizado un trabajo similar, con escáneres *Bluetooth* en su festival. En este caso se ha podido crear una base de datos con la que poder trabajar acerca de flujos de movilidad, modos de transporte utilizados y predicción de número de visitantes en los distintos momentos del festival (Versichele et al., 2012).

Finalmente, algunos trabajos han usado datos de GPS para analizar la vulnerabilidad a diferentes sucesos. En Wuhan (China) se ha construido una base de datos a partir de información obtenida de sistemas GPS del servicio de taxis para, mediante un matriz OD, analizar patrones de movilidad que circulan por tres puentes críticos que conectan los tres distritos de la ciudad, separados por el río Yangtsé. A partir de estos datos se identifican rutas alternativas que pasen por diferentes puentes y que observen las diferencias derivadas respecto a las trayectorias de la base de datos original (Fang et al., 2012). Para el tratamiento de la congestión, en Tokio, se está empleando un sistema de simulación a partir de navegadores de coches que permite estimar las condiciones de tráfico en tiempo real y medir niveles de ruidos de tráfico y de emisión de gases. El objetivo es diseñar un sistema de simulación de tráfico *nowcast* a partir de parámetros de calibración que incluyan volumen de tráfico y radio de saturación. La información se obtiene de navegadores GPS en automóviles y se calibra mediante un modelo matemático más preciso que las simulaciones convencionales (Hanabusa, 2012).

5.5. Información para conocer la percepción del servicio de transporte

Finalmente, el uso de *Big Data* basado en *Twitter* y la explotación de sus comentarios permiten conocer la valoración de los sistemas de transporte por parte de sus usuarios. Así, en la Universidad de California, ha buscado mejorar el servicio de tren en Los Ángeles a partir de opiniones públicas subidas en las redes sociales. Después de recopilar los mensajes, se han clasificado las cuentas de usuarios en diversos tipos (gobierno, servicios, agencias, personas, etc.), realizando a continuación un análisis exploratorio de relaciones entre *retweets* para poder identificar diversos temas. Después, se ha procedido con un análisis en el que se procesó el lenguaje empleado para escribir los mensajes en diferentes sentimientos, para así saber las opiniones más comunes e identificar porque una línea particular de transporte público tiene un sentimiento positivo o negativo (Luong y Houston, 2015).

6. Debilidades y retos de las nuevas fuentes de datos

El uso de nuevas fuentes de datos está cambiando el modo de aproximarnos al estudio de la movilidad urbana, pero todavía hay camino por recorrer para su consolidación como fuentes de información de movilidad. Es evidente que estas fuentes presentan inconvenientes, retos que hay que considerar y que provienen de la propia generación masiva de información. A diferencia de la información procedente de encuestas o censos, la naturaleza desestructurada de los datos implica un trabajo previo de depuración, homogeneización o preparación de los datos. Se trata de dar un salto desde los datos a la información (Bosque, 2015).

El principal reto consiste en organizar y gestionar toda la información que compone el *Big Data*, debido a que los datos provienen de fuentes dispersas y diferentes, en información no normalizada. En algunos casos los datos pueden ser irrelevantes o no tener veracidad, careciendo de estructura y metadatos, lo que reduce la precisión de la información. Por otro lado, el volumen de información generado es tan grande que las herramientas y técnicas usadas hasta ahora se muestran inadecuadas a la hora de procesar los datos. El volumen, variedad y velocidad de actualización de las bases de datos originadas a partir del *Big Data* exceden la capacidad de las tecnologías empleadas comúnmente (Shekhar et al., 2012). La necesidad de adquirir tecnología nueva y personal formado también puede conllevar un coste importante.

Otro problema radica en que estas fuentes suelen presentar ciertos sesgos, es decir, la información obtenida depende del tipo de población que la genera. El acceso a distintos dispositivos móviles, a navegadores, o incluso a redes sociales varía dependiendo de factores como la edad o el poder adquisitivo. Por lo tanto, es necesario contrastarlas y complementarlas con las fuentes de datos tradicionales. Tanto las fuentes tradicionales como las nuevas tienen sus propias ventajas e inconvenientes, y su combinación permite realizar un estudio rico en información, tanto cualitativa como cuantitativa y con una alta resolución espacial y temporal (Miralles-Guasch et al., 2015).

En otras ocasiones los problemas asociados pueden estar en la protección de la privacidad o de los derechos de propiedad intelectual. En muchos casos el acceso a los datos es difícil, al depender directamente de empresas privadas (es el caso de la obtención de datos de tarjetas bancarias). El caso concreto del uso de las redes sociales conlleva una serie de barreras. La línea entre la vida personal y la pública de los usuarios a veces es confusa y a la vez conlleva problemas a la hora de tratar la privacidad. Las propias trayectorias de movimiento espacio-temporales suelen ser repetitivas y predecibles, convirtiéndolas en potentes identificadores, huellas digitales en el mundo virtual (OECD, 2015). La dependencia de gráficos, videos y contenido autogenerado produce problemas en cuanto a la accesibilidad, y la propia seguridad de los datos puede estar expuesta por diversas ciberamenazas. En el análisis de la percepción, las empresas pueden mostrar preocupación hacia el criticismo público. Finalmente, conviene considerar que el panorama de las redes sociales puede cambiar en el futuro, pudiendo hacer a veces difícil la adaptación los nuevos cambios. Otros desafíos radican en la interoperabilidad con otras fuentes, el acceso para personas con discapacidades o cuestiones multiculturales (Bregman, 2012).

Finalmente, trabajar con todas estas nuevas fuentes de datos requiere seguir una serie de pasos para transformar esos datos en información con valor (Imagen 5). Este proceso suele empezar por la adquisición de datos en bruto, ya sea de medios analógicos como audios o videos, o digitales como datos GPS. Una vez obtenidos, se procesan e integran mediante técnicas de extracción, limpieza de información innecesaria, optimización, almacenamiento, metadatado, o fusión de múltiples fuentes. A continuación, con estos datos ya organizados se pueden identificar patrones, formular análisis, construir modelos, y visualizar resultados. Finalmente, estos resultados son interpretados obteniendo la información buscada (OECD, 2015).

Imagen 5: Pasos del tratamiento de los datos de *Big Data*

Fuente: Elaboración propia en base a OECD (2015)

En relación al desafío que conlleva el uso óptimo de los datos, un campo nuevo es el del enriquecimiento de la información geográfica, cuyo objetivo es aumentar la calidad y el conocimiento generado a partir de la combinación de distintas fuentes, modelos de entrada de datos, *data mining* exploratorio de redes sociales, testeo de hipótesis, etc. Destacan el enriquecimiento semántico que otorga mayor número de atributos, y el enriquecimiento geométrico que concede mayor escala y resolución, y mayor facilidad para visualización en 3D (Zipf, 2015).

En cuanto a la computación, algunos desafíos son la necesidad de formar equipos para el proceso de datos masivos, la visualización de los resultados, el requerimiento de un cambio en el cuadro de referencia que permita cambiar desde una perspectiva global a la de un individuo viajando a través de una red de transportes, el abaratamiento y mejora en la especificación de consultas frente a las rutas tradicionales, la creación de algoritmos flexibles que permitan trabajar con las distintas fuentes de datos, el uso de razonamiento geoespacial en detección e inferencias a través del espacio y del tiempo, el tratamiento de la privacidad de las fuentes y el uso del SBD para predicción de fenómenos (Shekhar et al., 2012).

Académicamente, hay preocupación por que se prime la ciencia cuantitativa sobre la cualitativa. La información que se genera está inevitablemente cargada con suposiciones particulares que sesgan el conocimiento. Para intentar remediarlo, se busca que la narrativa de las nuevas fuentes de datos asociadas a las *TIC* y al *Big Data* permita analizar el mundo de forma objetiva (Graham y Shelton, 2013).

Como cierre de este apartado, podemos apreciar las ventajas e inconvenientes del uso del *Big Data* cómo fuente en un ejemplo práctico. El Centro para Análisis Espacial Avanzado de la *University College* de Londres ha estado trabajando acerca del efecto de una tarjeta común de transporte público para el área metropolitana de Londres. La información de viajes de los distintos sistemas de transporte público ha sido extraída a partir de los datos de uso de dicha tarjeta. Gracias al uso de las *TIC*, de los datos generados, y del gran detalle de información, se ha podido hacer una aproximación que hubiera sido imposible de realizar utilizando las fuentes tradicionales. Sin embargo, la enorme cantidad de datos (1 billón de registros en un periodo de seis meses), el sesgo de la propia fuente (hay un porcentaje de viajeros que se quedan fuera del estudio al no usar la tarjeta de estudio, incluyendo grupos concretos como los turistas), o la dependencia de las propias infraestructuras de transporte (para que la información se registre, hace falta pasar la tarjeta por sensores en paradas; las estaciones con barreras abiertas conllevan un acceso al transporte sin necesidad de usar la tarjeta generándose vacíos de información en determinadas zonas que hay que sortear mediante estimaciones) conllevan una dificultad importante que hay que sortear para poder establecer análisis de movilidad (Batty, 2013).

7. Conclusiones

El desarrollo acelerado de las ciudades ha propiciado nuevos patrones de comportamiento en la movilidad tanto a nivel espacial como temporal, dándose como resultado un mayor número de viajes, más largos y dispersos. Ante este nuevo paradigma, es necesario realizar un estudio eficaz de la movilidad metropolitana. Sin embargo, las fuentes tradicionales se muestran ineficaces para monitorizar estos cambios debido a su baja resolución espacio-temporal.

En este panorama surgen las Tecnologías de la Información y Comunicación como mecanismos dinámicos y de alta resolución, que generan nueva información y abren un abanico nuevo de posibilidades en la investigación en ciencias sociales. Desde muy recientemente, empresas, organismos y universidades están empleando bases de datos basadas en el *Big Data* originadas a través de información transmitida por las TIC para poder investigar y resolver asuntos de diversos campos, siendo la movilidad una de las áreas donde se está dando un mayor uso de estas herramientas, gracias a que los datos pueden contar con información geográfica.

Uno de los factores que ha contribuido al *Big Data* es la tecnología móvil. Hoy casi toda la población de las grandes ciudades cuenta con un *smartphone* con acceso a Internet, lo que permite subir información de todo tipo desde cualquier espacio y georreferenciarla gracias a la arquitectura GPS. También es notorio señalar las tarjetas electrónicas inteligentes y los navegadores en vehículos como fuentes para el estudio de la movilidad.

La evolución de Internet y de la *Web 2.0* ha conllevado el auge de las redes sociales, programas de interacción donde se comparte información que se sube a la red. Algunas redes sociales cuentan con la posibilidad de compartir datos geolocalizados, siendo *Twitter* una de las más remarcables en este campo. Además, el tránsito a la *Web 3.0* conlleva el paso hacia la red del *Big Data*, donde gracias a las redes sociales prima la información abierta y accesible.

La capacidad de realizar análisis espaciotemporales a partir de bases de datos gigantes que contienen mensajes de redes sociales masivas como *Twitter* ha abierto nuevas ventanas a un proceso de investigación que busca aprovechar cada vez más las nuevas tecnologías para obtener resultados más reales y precisos, pudiendo realizarse a su vez nuevos tipos de análisis que facilitan hacer estudios demográficos o de movilidad a partir de grupos sociales, sentimientos o palabras. Cada vez hay mayor número de aplicaciones destinadas al estudio de la movilidad como la creación de *datacubes*, la extracción de matrices origen-destino, o la creación de cartografía interactiva, destinados a objetivos como el estudio de la accesibilidad, la gestión de eventos, la identificación de espacios con mayor tránsito, o la mejora del transporte.

En el campo de la movilidad urbana, el objetivo de transformar el modo de trabajo de los modelos de transporte es esencial para dar un punto de inicio al entendimiento de las ciudades como sistemas complejos con su propio comportamiento (Serras, Bosredon, Herranz, y Batty, 2014). El *Spatial Big Data* puede ser usado como herramienta para conseguir un transporte más satisfactorio y sostenible gracias a su transparencia y flexibilidad (Baker y Thomas, 2012). Además, gracias a los nuevos datos es posible identificar dinámicas a medio o largo plazo en contraposición con las bases de datos tradicionales que están adecuadas principalmente al corto plazo (Long y Shen, 2015).

Para concluir, el campo del *Big Data* para el estudio de la movilidad todavía está por explorarse y desarrollarse. Los principales retos de cara al futuro consisten en el enriquecimiento de datos,

la mejora de la precisión, la credibilidad y el contraste de información, la interoperabilidad entre distintas fuentes, el metadato, la combinación de diferentes formatos, la detección de patrones en datos desestructurados, la protección de la privacidad, la seguridad, o el uso de filtros y sesgos.

Agradecimientos

Los autores agradecen la financiación recibida de la Comunidad de Madrid (SOCIALBIG-DATA-CM, S2015/HUM-3427), del Ministerio de Educación, Cultura y Deporte (Programa FPUAP2015-0147), y del Ministerio de Economía y Competitividad y el Fondo Europeo de Desarrollo Regional (FEDER) (Proyecto DynAccess, TRA2015-65283-R).

8. Referencias bibliográficas

- Baker, R., y Thomas, L. (2012). “Modelling real people and creating better place and movement plans: irrationality, Big Data and increasing access to choice”. *LTT’s Modelling World 2012*.
- Badger, E. (2012). “You already own the most important transportation planning tool”. <http://www.citylab.com/tech/2012/02/you-already-own-next-most-important-transportation-planning-tool/1124/> [Consulta: 20 de octubre de 2016].
- Banister, D. (2008). “The sustainable mobility paradigm”. *Transport Policy*, 15(2), 73–80.
- Banister, D. (2011). “Cities, mobility and climate change”. *Journal of Transport Geography*, 19(6), 1538–1546.
- Batty, M. (2013). “Big data, smart cities and city planning”. *Dialogues in Human Geography*, 3(3), 274–279.
- Birkin, M., Harland, K., Malleson, N., Cross, P., y Clarke, M. (2014). “An Examination of Personal Mobility Patterns in Space and Time Using Twitter”. *International Journal of Agricultural and Environmental Information Systems*, 5(3), 55–72.
- Blanford, J. I., Huang, Z., Savelyev, A., y MacEachren, A. M. (2015). “Geo-located tweets. Enhancing mobility maps and capturing cross-border movement”. *PLoS ONE*, 10(6), 1–16.
- Bosque, J. (2015). “Neogeografía , Big Data y Tig : Problemas y nuevas posibilidades”. *Polígonos*, 27(2007), 165–173.
- Bregman, S. (2012). “TCRP Synthesis 99: Uses of Social Media in Public Transportation”. http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_syn_99.pdf [Consulta: 18 de septiembre de 2016]
- Caceres, N., Romero, L. M., Benitez, F. G., y del Castillo, J. M. (2012). “Traffic Flow Estimation Models Using Cellular Phone Data”. *IEEE Transactions on Intelligent Transportation Systems*, 13(3), 1430–1441.
- Cao, G., Wang, S., Hwang, M., Padmanabhan, A., Zhang, Z., y Soltani, K. (2014). “A Scalable Framework for Spatiotemporal Analysis of Location-based Social Media Data”. *Computers, Environment and Urban Systems*, 51, 70–82.
- Ciuccarelli, P., Lupi, G., y Simeone, L. (2014). “Visualizing the Data City Social Media as a Source of Knowledge for Urban Planning and Management”. <https://doi.org/10.1007/978-3-319-02195-9>
- Clarke, M. (2016). “Big Data in Transport”. *Institution of Engineering and Technology Sectors Insights*, 1–70.
- Dewulf, B., Neutens, T., Vanlommel, M., Logghe, S., De Maeyer, P., Witlox, F., ... Van de Weghe, N. (2015). “Examining commuting patterns using Floating Car Data and circular statistics: Exploring the use of new methods and visualizations to study travel times”. *Journal of Transport Geography*, 48, 41–51.
- Europe, E. (2015). “Big Data Europe for Smart , Green and Integrated Transport 1st Workshop Report”. *22th World Congress on ITS*.
- Fang, Z., Shaw, S. L., Tu, W., Li, Q., y Li, Y. (2012). “Spatiotemporal analysis of critical transportation links based on time geographic concepts: A case study of critical bridges in Wuhan, China”. *Journal of Transport Geography*, 23, 44–59.
- Finnis K, K., y Walton, D. (2007). “Field observations of factors influencing walking speeds”. *International Conference on Sustainability Engineering and Science, 2Nd, 2007, Auckland, New Zealand*, 13P.
- Frias-martinez, V., Soto, V., Hohwald, H., y Frias-martinez, E. (2012). “Characterizing Urban Landscapes using Geolocated Tweets”. *2012 International Conference on Social Computing*.

- Gao, S., Yang, J., Yan, B., Hu, Y., Janowicz, K., y McKenzie, G. (2014). "Detecting Origin-Destination Mobility Flows From Geotagged Tweets in Greater Los Angeles Area". *Eight International Conference on Geographic Information Science*, 1–5.
- García-Palomares, J. C. (2010). "Urban sprawl and travel to work: the case of the metropolitan area of Madrid". *Journal of Transport Geography*, 18(2), 197–213.
- García-Palomares, J. C., Gutiérrez, J., y Mínguez, C. (2015). "Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using photo-sharing services and GIS". *Applied Geography*, 63, 408–417.
- Gasparini, A., y Guidicini, P. (1990). *Innovazione tecnologica e nuovo ordine urbano*. Milán: Angeli.
- GHD. (2014). "New traffic data sources – An overview". *New Data Sources for Transport Workshop, BITRE*, 29.
- Graham, M., y Shelton, T. (2013). "Geography and the future of big data, big data and the future of geography". *Dialogues in Human Geography*, 3(3), 255–261.
- Griffiths, R., y Richardson, A. J. (2000). "Travel Surveys". *Transportation in the New Millenium*.
- Gutiérrez-Puebla, J., García-Palomares, J. C., y Salas-Olmedo, M. H. (2016). "Big (Geo) Data en Ciencias Sociales: Retos y Oportunidades". *Revista de Estudios Andaluces*, 33(331), 1–23.
- Hanabusa, H. (2012). "Development of Nowcast Traffic Simulation System for Road Traffic in Urban Area". *20th World Congress on ITS*, 10, 3–10.
- Hasan, S., Zhan, X., y Ukkusuri, S. V. (2013). "Understanding Urban Human Activity and Mobility Patterns Using Large-scale Location-based Data from Online Social Media". *Proceedings of the 2Nd ACM SIGKDD International Workshop on Urban Computing*, 6:1--6:8.
- Kitchin, R. (2013). "Big data and human geography: Opportunities, challenges and risks". *Dialogues in Human Geography*, 3(3), 262–267.
- Kwan, M., y Lee, J. (2011). "Visualisation of Socio Spatial Isolation Based on Human Activity Patterns and Social Networks in Space Time". *Tijdschrift Voor Economische En Sociale Geografie*.
- Lansley, G., y Longley, P. A. (2016). "The geography of Twitter topics in London". *Computers, Environment and Urban Systems*, 58, 85–96.
- Lathia, N., Smith, C., Froehlich, J., y Capra, L. (2013). "Individuals among commuters: Building personalised transport information services from fare collection systems". *Pervasive and Mobile Computing*, 9(5), 643–664.
- Lee, J. H., y Lee, J. H. (2015). "Can Twitter data be used to validate travel demand models?". *95th Annual Transportation Research Board Meeting*, 1–27.
- Lenormand, M., Louail, T., Cantú-Ros, O. G., Picornell, M., Herranz, R., Arias, J. M., ... Ramasco, J. J. (2015). "Influence of sociodemographics on human mobility". *Scientific Reports*, 5(8557), 10075.
- Lenormand, M., Picornell, M., Cantú-Ros, O. G., Tugores, A., Louail, T., Herranz, R., ... Ramasco, J. J. (2014). "Cross-checking different sources of mobility information". *PLoS ONE*, 9(8), 30–38.
- Long, Y., y Shen, Z. (2015). "Profiling Underprivileged Residents with Mid-term Public Transit Smartcard Data of Beijing". *Geospatial Analysis to Support Urban Planning in Beijing* (pp. 169–192). Springer International Publishing.
- Longley, P. A., y Adnan, M. (2016). "Geo-temporal Twitter demographics". *International Journal of Geographical Information Science*, 30(2), 369–389.
- Luo, F., Cao, G., Mulligan, K., y Li, X. (2016). "Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago". *Applied Geography*, 70, 11–25.
- Luong, T. T. B., y Houston, D. (2015). "Public opinions of light rail service in Los Angeles , an analysis using Twitter data". *iConference 2015 Proceedings*, 2–5.
- Martin, D. J., Jordan, H., y Roderick, P. (2008). "Taking the bus: Incorporating public transport timetable data into health care accessibility modelling". *Environment and Planning A*, 40(10), 2510–2525.
- McKinsey & Company. (2011). "Big data: The next frontier for innovation, competition, and productivity". *McKinsey Global Institute*, (June), 156.
- Miralles-Guasch, C. (2012). "Las encuestas de movilidad y los referentes ambientales de los transportes". *Eure*, 38(115), 33–45.
- Miralles-Guasch, C., Delclòs, X., y Vich, G. (2015). "Nuevas fuentes de información para el análisis de la movilidad cotidiana: de las encuestas de movilidad a las aplicaciones para móviles". *XXIV Congreso de La Asociación de Geógrafos Españoles*, 2055–2063.

- Miralles-Guasch, C., y Martínez, M. (2013). “Las fuentes de información sobre movilidad: la visión de los profesionales. Ejemplo de aplicación de metodología DELPHI”. *Revista Transporte Y Territorio*, (8), 99–116.
- Moro, E. (2014). “Tweeting and moving around. A day of trips in Spain”. <https://vimeo.com/111579945> [Consulta: 23 de octubre de 2016].
- Moro, E. (2016). “Ciudades Movilidad y Social Media”. *VII Congreso Estatal RITSI*.
- Netto, V. M., Pinheiro, M., Meirelles, J. V., y Leite, H. (2015). “Digital footprints in the cityscape: Finding networks of segregation through Big Data”. *International Conference on Location-Based Social Media Data*, 1–15.
- OECD. (2015). “Big Data and Transport: Understanding and assessing options”. *International Transport Forum*.
- Ortúzar S., J. de D., y Willumsen, L. G. (2011). *Modelling transport*. Oxford: Wiley-Blackwell.
- Pazos, M. (2005). “El estudio de la movilidad diaria en España: limitaciones en las fuentes y alternativas propuestas”. *Eria*, 66, 85–92.
- Pereira, F. C., Rodrigues, F., y Ben-Akiva, M. (2015). “Using Data From the Web to Predict Public Transport Arrivals Under Special Events Scenarios”. *Journal of Intelligent Transportation Systems*, 19(3), 273–288.
- Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., y Srivastava, M. (2010). “Using mobile phones to determine transportation modes”. *ACM Transactions on Sensor Networks*, 6(2), 1–27.
- Romanillos, G., y Zaltz Austwick, M. (2015). “Madrid cycle track: visualizing the cyclable city”. *Journal of Maps*, 12(5), 1–9.
- Sareh, H. A., Mohammad Ali, N., y Khosravi Farsani. (2012). “Evolution of the World Wide Web : From Web 1.0 to Web 4.0”. *International Journal of Web & Semantic Technology*, 3(1), 1–10.
- Schwanen, T. (2016). “Geographies of transport II: Reconciling the general and the particular”. *Progress in Human Geography*, 1–10.
- Serras, J., Bosredon, M., Herranz, R., y Batty, M. (2014). “Urban Planning and Big Data - Taking LUi Models to the Next Level?” <http://www.nordregio.se/en/Metameny/Nordregio-News/2014/Planning-Tools-for-Urban-Sustainability/Reflection/> [Consulta: 11 de octubre de 2016].
- Shekhar, S., Evans, M. R., Gunturi, V., y Yang, K. (2012). “Spatial big-data challenges intersecting mobility and cloud computing”. *MobiDE 2012 - Proceedings of the 11th ACM International Workshop on Data Engineering for Wireless and Mobile Access - In Conjunction with ACM SIGMOD / PODS 2012*, 1(c), 1–6.
- Shelton, T., Poorthuis, A., y Zook, M. (2015). “Social media and the city: Rethinking urban socio-spatial inequality using user-generated geographic information”. *Landscape and Urban Planning*, 142, 198–211.
- Steiger, E., Westerholt, R., Resch, B., y Zipf, A. (2015). “Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data”. *Computers, Environment and Urban Systems*, 54, 255–265.
- Steur, R. (2014). “Twitter as a spatio-temporal information source for traffic incident management”. *Geographical Information Management and Applications*.
- Sui, D., y Goodchild, M. (2011). “The convergence of GIS and social media: challenges for GIScience”. *International Journal of Geographical Information Science*, 25(11), 1737–1748.
- Tao, S., Rohde, D., y Corcoran, J. (2014). “Examining the spatial-temporal dynamics of bus passenger travel behaviour using smart card data and the flow-comap”. *Journal of Transport Geography*, 41(December), 21–36.
- Tascón, M., y Coullaut, A. (2016). *Big Data y el internet de las cosas : qué hay detrás y cómo nos va a cambiar*. Madrid: Catarata.
- Versichele, M., Neutens, T., Delafontaine, M., y Van de Weghe, N. (2012). “The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities”. *Applied Geography*, 32(2), 208–220.
- Wu, L., Zhi, Y., Sui, Z., y Liu, Y. (2014). “Intra-urban human mobility and activity transition: Evidence from social media check-in data”. *PLoS ONE*, 9(5).
- Zhao, F., Ghorpade, A., Pereira, F. C., Zegras, C., y Ben-Akiva, M. (2015). “Quantifying Mobility: Pervasive Technologies for Transport Modeling”. *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, 1039–1044.
- Zipf, A. (2015). “Enrichment of volunteered geographic information: some considerations”. *RICH-VGI: Enrichment of Volunteered Geographic Information (VGI): Techniques, Practices and Current State of Knowledge*, 1–15.

Sobre los autores/as

JOAQUÍN OSORIO ARJONA

Estudiante de Doctorado en Geografía en la Universidad Complutense de Madrid con beca de financiación FPU por el Ministerio de Educación, Cultura y Deporte (AP2015-0147), con Máster Oficial en Tecnologías de la Información Geográfica por la Universidad de Alcalá. Miembro del grupo de investigación t-GIS: Transporte, Infraestructura, y Territorio. Actualmente está desarrollando la tesis doctoral “*Spatial Big Data y movilidad en espacios urbanos a partir de TICs*”. Sus líneas de investigación se basan en la cartografía aplicada al transporte y urbanismo, y el estudio y diseño de matrices origen-destino de movilidad a partir de nuevas tecnologías de la información y comunicación.

JUAN CARLOS GARCÍA PALOMARES

Juan Carlos García Palomares es Profesor Titular y miembro del Grupo de Investigación tGIS (Transporte, Infraestructura, y Territorio) en la Universidad Complutense de Madrid. Sus líneas de investigación son el transporte, la movilidad y la aplicación de los Sistemas de Información Geográfica (GIS).