The democratisation of data: a challenge for transport studies?

By
Catherine Morency (Ingénieure)
02 September 2013

Will the technological-IT revolution help us to better understand transport problems? Catherine Morency investigates.

Over the last few years, a number of new devices have been developed that have completely changed the way we access information. What I'm going to talk about here is precisely this revolution as regards access to information and data, and how it has changed practices in terms of the study of mobility studies. The era when people relied only on paper maps is disappearing. For the younger generation, it's now very easy to use different tools and devices to find out where you are and, more importantly, what route to take to reach different destinations in a single click. In addition to its impact on people's daily lives, these technologies and the information they give us access have affected and continue to affect the way transport studies are conducted. If we go back in the history of transport models, the first ones were developed in the 1950s. Early computers made it possible to compile certain data and make somewhat complex calculations; it was obviously a great simplification of the reality, and therefore something of a minor revolution.

Data and technology: towards a minor revolution

Today, the world of transport is experiencing a revolution; a revolution of access and the tools with which it can conduct studies, due mainly to the proliferation and democratisation of data and the technologies required to manage it. On the data side, one automatically thinks of Google and the various products it offers - Google Maps, Google Street View, Google Traffic, Google Earth, etc. Some of these have really changed the way people perceive their environment and understand streets and maps. There were predecessors and competitors, such as MapQuest and Yahoo Maps, and of course all of the features that went with them: maps, the ability to see your exact location, and several ways of making calculations and creating itineraries. Many of the features embedded in these tools have not only democratised access to data, but also the ease with which you can make calculations. In terms of technology, one thinks of mobile phones, tablets, GPS-equipped computers and even accelerometers. Together, these tools increase the possibilities for collecting data and are obviously all web-enabled. Thus, people can be connected and access information, which also ensures continuous sharing.

Towards a democratisation of data

Today, people and institutions have easy access to geographical information as well as information about transport networks. Individuals obviously change their behaviour as a result of having access to this information, about traffic congestion, possible routes, in other words, it slightly changes their way of thinking because they are now aware of the different alternatives. Individuals also create data during their journeys and can therefore themselves become contributors to this data that is so useful for the transport field. In the past, institutions had to invest a great deal in this area. A few years ago, access to data on regions, addresses and intersections was extremely expensive; oft times we had to create the data ourselves for it to be fit with transport models. In short, a lot of human resources and money were spent keeping databases up to date. Today, all of this can be delegated, there are numerous companies that sell databases or simply make them available online. This really is democratisation in action. Nowadays, lots of people have access to data, a better understanding of what it means and greater awareness of its value. Obviously, the information is "monitored". The more users we have of these tools, the better we can guarantee a certain level of quality. If someone doing a Google search notices a street is missing, they can send a message or a comment to Google so that the information can be updated. We therefore have a series of "monitoring agents" who help to improve the quality of the data, which is also a very good technique.

Data that's more recent and more complete, and that costs less

How does all this affect transport planning and studies? In terms of data collection, the first step is to obtain the data to conduct the studies. The emergence of these tools like the Internet and Google, with all their databases on transport networks and regions, has made it easier to access this information. Tools can be developed that enable us to collect data from the web about the mobility of dwellers. In several regions, major "household surveys" are carried out, usually by telephone or face-to-face. The availability of all this data and the fact that it can be accessed via the web makes it possible to reach people by way of this new technology. Indeed, we see that the data streaming now available via chips and smart cards, like data from users of public transport, will certainly shake up our habits in terms of data collection. Often, after setting up a data collection process, once again, at considerable expense in terms of human and financial resources, we obtain the actual results only a year later, after the operation is complete, by which time it's virtually out of date. So the web, and the availability of platforms like those offered by Google, which combine reference data, files and calculating functions, facilitate the development of these collection tools. Another interesting feature is that they can be transposed more easily from one region to another, because Google offers far more comprehensive coverage than what's usually available for a given area.

Information that can be used directly by users, transport authorities… and Google!

Another interesting aspect in the treatment of this data is the fact that such features are more easily accessible to average developers, meaning that someone can develop a study using Google features in the comfort of their own home. The other element that has had a major impact on the transport world is the emergence of standards. Here I'm talking about GTFS files, which stands for General Transit Feed Specification. These are the files used to describe public transport networks. Several cities have adopted them, which has enabled Google to introduce trip calculators for public transport networks. The beauty of all this is that the development of features based on these standard files can be easily transposed from one region to another. So any region that adopts this standard can then use features that have been developed all around the world, thus facilitating comparative analyses and allowing for calculators for different types of transport networks: road, public transport, etc. It's true that this takes the responsibility off the shoulders of transport authorities, that is, it removes the burden of developing their own journey planning tools, but that's not such a bad thing. The fact that these tools are available in different regions also makes them more accessible to transport users: the information is better and some applications can be downloaded onto mobile phones. Obviously, as I said earlier, Google Maps

and its predecessors MapQuest and Yahoo Maps revolutionised the field by facilitating access to data that previously was cumbersome to develop and quickly became obsolete because of the sheer cost of continuously updating it. Since then, they have developed features that allow people to access information that previously was impossible to find. Of course, every good thing has its downside. Google is still a company, and you cannot necessarily use the data or tools as you like. Google maintains a degree of control over the way this data is used, and we can clearly see the ethical ambiguity in this. That is, if you open your phone and do an itinerary search, Google will ask for permission to use your location data. If you answer 'yes', you effectively contribute to a data set that allows the company, in this case Google, to develop tools or evaluate traffic congestion. Google Traffic, for example, which was put online a few years ago, is based on the GPS traces from trips made by people who authorised Google to use their data. Obviously, as a Google representative would explained it, when you decide to launch Google Maps using the "My location" option on your mobile phone or tablet, the device sends out anonymous tracking data about your speed and your spatial location. When we compile all the data from all of the people travelling, we estimate journey times, and hence traffic congestion. Naturally, it's not obvious who owns that data; neither is it obvious that people who authorise Google to monitor their movements are aware they are not only helping to provide very useful tools, but are also contributing to the use that Google can make of that data and the profits the company can generate as a result!

Open and credible alternatives to Google

Not surprisingly, we have also witnessed the emergence of other types of 'open' tools, such as Open Street Map, which are in the public domain. It is a website that relies on the contributions of almost everyone. I'd say it is what Wikipedia is to the encyclopaedia world, but instead providing geographical information. Based on these open sources, new developments and features are being created that, while similar to those available on sites like Google, are based on open-source data and can be used without any licensing. Obviously, you might think these open sites would cause chaos or that their data wasn't fully representative of reality, but that's not what we're seeing. Wikipedia is proof of this; the content is fairly representative and there are even studies that show that it's roughly the equivalent of what you would find in the Encyclopedia Britannica. So, by drawing on each person's individual contribution, a technique known as crowdsourcing, it's possible to tap into the creativity, intelligence and expertise of several people, who can be either hired or can work for free to create data sets that will be useful for whoever wants to use them.

A challenge for transport studies

So, this too is a revolution. We have seen the emergence of major websites with a multitude of features that have enabled many people to access data that was once very expensive to develop. This data is now widely available: everyone has access to it and everyone is free to use it for their own purposes. The challenge now is no longer access, which was once a hurdle, but no longer is, but using it in a meaningful way. The amount of data available has increased, which obviously helps us to have a better understanding of the complexity of behaviours. If you have data on every single day trips being made, it's possible to properly understand variability, have up-to-date information and know when new roads are being built, but this obviously involves a lot data handling and paves the way for more complex models. Therefore, a lot of work will have to be done to improve the models that use this data as well. Indeed, if we have a data streaming and a deeper understanding of complexity, we shall take this into account when developing new transport models. Obviously, the democratisation of data, combined with the democratisation of tools, has paved the way for new methods and models that are much more complex and collaborative. And so now, we can call on people to contribute to the quality of the data and question them in a simpler manner. What's more, another aspect, visualisation, or the way this data is disseminated, is also being developed based on this open data and the features available in Google

and tools like Open Street Map make it possible to develop new objects that will enable more people to understand all of these transport issues.

Mobility

For the Mobile Lives Forum, mobility is understood as the process of how individuals travel across distances in order to deploy through time and space the activities that make up their lifestyles. These travel practices are embedded in socio-technical systems, produced by transport and communication industries and techniques, and by normative discourses on these practices, with considerable social, environmental and spatial impacts.

En savoir plus x

Associated Thematics :

Lifestyles

- Digital technologies

Policies

- Cars
- Aviation
- Public transport
- Cycling & Walking
- Cities & Territories

Theories

- Methods

Catherine Morency

Ingénieure

Catherine Morency is a civil engineer and professor of transportation planning at Polytechnique Montréal. She heads the Mobilité Chair which focuses on the assessment and implementation of sustainable transportation and the Canada Research Chair which studies interactions between different modes of transportation. Her research concentrates on modeling individual travel behaviors

including active and alternative modes of transportation, data collection methods and the development of simulation tools.

The same author See all publications
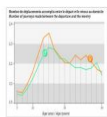
Anticipating transport needs based on a comparative analysis of men's and women's mobility

Video by Catherine Morency

Sustainable Mobility: definitions, concepts and indicators

Video by Catherine Morency

To cite this publication :

Catherine Morency (02 September 2013), « The democratisation of data: a challenge for transport studies? », Préparer la transition mobilitaire. Consulté le 12 May 2024, URL: https://forumviesmobiles.org/en/videos/1137/democratisation-data-challenge-transport-studies
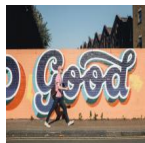

Other publications

How to develop territories to promote more localized lifestyles?

Master 2 et Magistère d'Urbanisme et Aménagement Université Paris 1 (2021-2022)

In the post-Covid world, have we forgotten about walking again?

Jean-Marc Offner

"We need to boost public transport supply and limit car use in order to reduce our emissions by 60% within 10 years"

Jean Coldefy

Are homeless people "vagrants"?

Projet collectif SciencesPo/Architecture Bordeaux

1 http://creativecommons.org/licenses/by-nc-sa/3.0/fr/
2 http://forumviesmobiles.org
3 http://creativecommons.org/licenses/by-nc-sa/3.0/fr/
4 http://fr.fvm.localhost/modal_forms/nojs/contact