Demystifying big data

by Maxime Flament, *Head of Connected and Automated Driving,*
and Andrea Toth, *Communications Officer, ERTICO-ITS Europe*

Big data tends to be looked upon as the silver bullet for a handful of industries. But there's a big problem with so much attention and publicity, namely the inevitable pressure that comes from the need to live up to it. This is a pity because big data can, in fact, help a lot. It's maybe not a magic weapon, yet has all the potential to be a good tool in better understanding the complex world we're living in.

There are plenty of ways we can describe the 21st century's economy – from clichés like the globalised economy or the sharing economy, to more exciting phrases like the attention economy. But perhaps one of the most fitting expressions is the information economy. It hints to multiple aspects: A shift from traditional industries to ICT-based commerce, attaching value to information, an enabler for better decisions, and finally, on a more microeconomics base, it is the fuel driving a competitive edge among companies.

Four Vs fuelling the information economy

Information derives from data, therefore the better quality of the data we gather, the better quality of the information and knowledge we can potentially possess. However, it's not solely the volume of data that will ensure its quality; it's also the veracity (levels of uncertainty of reliability), velocity (speed of flow and processing), and variety (different forms and sources). The combination of these four Vs is what ensures that the information we acquire is accurate, of high quality, and can serve as a good enough base for analysis. Next, it's up to us to make better decisions based on the knowledge we have extracted so meticulously.

Big data technologies are nothing more than tools making it possible to collect, synchronise, interpret, analyse, and visualise all the data we can gather. It may even seem out of date to call them big data because what we're truly searching for are smart data.

We've never had so much data available as we do today, whether it is in our personal or professional life, as a company or a city, or an intergovernmental organisation, the amount of data humanity as a whole is generating is absolutely colossal. Perhaps the most mind-blowing fact comes from a 2013 study, quoted by most journalists, marketers, and data scientists, stating that 90% of all the world's data was created in the past two years. That's a lot of information, which is great at first glance, but if we cannot make sense of it, it's pretty much useless.

And here come big data technologies, which are nothing more than tools making it possible to collect, synchronise, interpret, analyse, and visualise all the data we can gather. It may even seem out of date to call them big data because what we're truly searching for is smart data.

Big data in big (and small) cities

By nature, public authorities usually lag a bit behind commercial businesses when it comes to adopting innovative technologies and new business models. But as cities are facing the pressure of a growing urban population, it seems inevitable to look towards novel ways of managing overpopulation, congestion, and pollution. Whether the solution lies in town

planning, improving public transport services, or envisioning fleets of driverless vehicles (read BTJ 5/15's *Reinventing urban mobility. Self-driving car fleets*), they all require large amounts of real-time, accurate, and accessible data.

Smart cities are one of those vaguely defined concepts that we keep hearing about incessantly. In fact, you may not even know it and you're living in one. At their very core, smart cities utilise the power of data and connectivity to enable a better functioning environment. Like managing traffic lights in line with the street's flow, or introducing a smart parking system.

Nevertheless, city-managing mobility is a tough one. The larger the city, the bigger the problem. Though restricting access, such as introducing traffic fees, could have its merits, a future solution may want to focus less on private cars and more on other players in the traffic system. The road, the buildings, the busses and taxi fleets are all part of the same city, the same problem, and could potentially be part of the solution. Thousands of sensors are constantly recording massive amounts of data which can help not only in predicting traffic trends, but also with making emergency events more seamless.

Having better functioning public transport can contribute to decreasing congestion and better traffic management on its own. However, in too many cities, current solutions for public transport are "just not good enough" to substitute a car. It isn't very easy to appeal to the better side of people: Yes, we recycle, and sure, if we like to, we bike around, but honestly, cars are still more convenient in comparison to other modes. Using big data applications could potentially improve the public transport system with minimal investment. It could help with balancing supply and demand such as by tracking passenger flow and managing the fleet accordingly. Or provide customized solutions to users, and incentivize or penalize the use of a specific route choice as opposed to another.

Before you can map out trends and advise, you first need to gather, validate, link, and store the data, which is an ongoing challenge for many cities with a limited budget. But big data solutions don't have to be expensive at all, while their specific return-on-investment can considerably outweigh the initial set-up capital and operational expenses required, including time (and traffic jam stress) savings for citizens, decreased road-caused pollution hence health and environmental benefits (read about trillions of US dollars of mortality and morbidity costs from ambient air pollution in BTJ 4/14's *Driving out of last breath. Cost of air pollution from road transport*), better logistics, public money set free for other investments (e.g. green areas instead of new traffic lanes or parking lots), and so on and so forth. In this regard, the EU-funded Big Data Europe project has recently unveiled its open source platform to tackle challenges related to big data in the transport domain. The project has a wide scope in tackling issues relating to seven societal challenges, the one regarding transport focuses on traffic management in cities.

My car my data?

Whether it is parking, traffic management, public transport enhancement, or another way of improving urban mobility, the options all fit into one giant ecosystem. You can always develop bits of the puzzle and bring new actors in on the fly.

Connecting all these parts is where, among others, the Internet of Things (IoT) comes in (read more in BTJ 1/16's *Connecting the unconnected. The Internet of Things in warehousing and logistics*). Motion sensors fitted on a delivery bike, a tram, or a piece of clothing transmit so much data that the dynamics of a city can be mapped out. The more dots we connect, the more accurate we become, the greater the potential to serve a public good. And there are plenty of connected things available already – they just need to start talking to each other.

Data privacy is a hot topic in whichever industry you tap into. When you look at transport it seems obvious to look at navigation or speed, but it's not limited to what the car provides on its own – it's also data transmitted via your smartphone, your online posts, your tweets, etc. There's been an unresolved debate for a long time on ownership of data shared online, and regulation has an important role in defining how to deal with privacy in the first place. Even defining who owns the data (i.e. whose privacy needs to be protected) is not as straightforward as it may seem.

Getting access to sufficient amounts of data for the big data effect to kick in is important and restricting access, in some cases, does more harm than good. In that sense, the type of data benefiting the society can be looked upon as an enabler for a greater public good.

At the same time, getting access to sufficient amounts of data for the big data effect to kick in is important and restricting access, in some cases, does more harm than good. In that sense, the type of data benefiting the society can be looked upon as an enabler for a greater public good.

Clearly, it's controversial, and this is not the only industry where legislation quickly becomes obsolete, not being able to keep up with the technological advancements. Laying down a set of key principles and processes could be more beneficial than defining specific instances (which could become pointless in a year or two).

Disrupting the world day by day

Data production will likely not slow down: Being connected, being mobile, being a data producer and user simultaneously is the reality today. There are always other competitors though in the runner up for "the next magic disruptor," and many of them tend to dismiss big data as a buzz word or a momentary fad.

Maybe it is, and maybe it has been hyped up too much and now has to live up to revolutionary qualities impacting our everyday lives. In reality, revolutions don't come so often, and it can make us cautious when we run into headlines of big data every other day, or even worse, bored and immune (e.g. virtual reality was such big excitement in the mid-1990s, and a total disaster at that time, which required another two decades to come to the surface again). Maybe we should let it be what it is: A great technology with the promise of helping us to make decisions a bit easier, a bit faster, and a bit better. In the end, this "bit" can deliver a big difference. ■

ERTICO – ITS Europe is a public/private partnership of over 100 companies and organisations across eight different sectors advocating, and actively working on, the deployment of Intelligent Transport Systems and Services (ITS) through a wide range of activities, such as events, projects, and interest groups. For more info, please visit http://ertico.com.

Big Data Europe is a project within the EU's Horizon 2020 framework programme (grant agreement No. 644564), which aims at building a knowledge- and innovation-based society, to strengthen the competitiveness of Europe's economy. For more info, please visit www.big-data-europe.eu.