

POLIS POSITION PAPER

THE MOVE TOWARDS OPEN DATA IN THE LOCAL TRANSPORT DOMAIN

JUNE 2013

A. Local authorities and transport data

1. Most transport data held/owned by local (transport) authorities is gathered/used primarily for **network management purposes**. Traffic regulations, such as speed limits, one-way streets or access restrictions, and real-time traffic or travel data, are first and foremost tools to support the transport authority manage the transport network in a safe, efficient and sustainable manner.
2. While **information services** have traditionally built upon the vast amounts of static and dynamic transport data available, recent years have seen a growth in data collected specifically for the information services market. Nonetheless, the vast majority of data held by local authorities remains first and foremost a tool for managing traffic, public transport and other mobility services.
3. **Data is not always used**, or even stored – in particular, real-time data is generally discarded, unless there is specific research that requires its storage. System performance data (e.g. fault monitoring, signal settings, etc.) would be retained only for a limited period
4. The **quality** of data within local authorities can be highly variable:
 - i. Most local authorities monitor only a fraction of the total network
 - ii. Data can be inconsistent in quality and completeness
 - iii. Manually-acquired data generally requires “cleaning”, and would be more prone to misinterpretation by others
5. The multiple tasks of gathering, processing and storing data comes at a significant **cost** for transport authorities. While the data may not be fully reliable, it is generally sufficient for network management purposes.
6. Within any given administrative area, **transport data tends to be widely dispersed** among the different transport bodies (public or private) and indeed within any single body. In larger areas, the transport data picture may be very complex and there is probably no clear overview of what transport data is available.
7. Due to the great number of contracts between public and private bodies for the provision of transport (-related) services, the **data ownership** situation is not always clear. This was confirmed in a recent Polis survey which showed that both traffic and public transport data ownership is shared between the public and private sector.
8. The move to open data is forcing the public sector to review ownership and transmission issues in their contracts with third parties. Specifically, the publication or release of data that may impact on

legal or contractual proceedings involving third parties is a source of concern, and precludes the widespread release of data in the first instance.

B. Local authorities and open data

1. In the last few years, the open data movement in the transport domain has taken off. Local authorities are increasingly moving towards open data as part of a wider public sector transparency agenda. Open transport data is rarely a legal requirement but is rather policy driven. Caution should therefore be exercised when adopting legislation: regulation could serve to stifle innovation in the fast evolving area of open data, and could even cause some existing open data initiatives to shut down.
2. In addition to transparency, open data offers local authorities, an opportunity to meet other **local transport policies**, notably to promote sustainable travel choices, by enabling them to:
 - i. relook at their own business and **improve internal processes**, notably (i) seeking to understand which data a public body holds/gathers, (ii) thinking strategically about the value of data, (iii) improving the quality of data through feedback from the developer community.
 - ii. improve the **quality of service** to users by harnessing the creativity of the apps developer community to produce innovative services based on one or more data feeds.
 - iii. reduce the **cost of service provision**, by allowing the local authority to focus on data acquisition and management while the private sector takes on some of the burden of disseminating information to users.
 - iv. promote **economic development**, especially for local information services providers (see previous point).
3. Experiences to date have revealed a number of local **challenges** in making transport data available. They include:
 - i. **Opposition from information service providers**, due in some cases to the fear of the threat of competition.
 - ii. **Data control and ownership**, for instance where data is owned by a cross-agency institution (eg, passenger transport authority) or data is provided by a contractor. Some concerns cited by contractors, especially those operating in the passenger transport sector, include: 'competitor or commercially sensitive', 'fear of use for measuring operational performance', 'extra burden on operations'.
 - iii. Organisational: there is **no clear process/practical framework** to guide transport authorities in opening up their data. Open data is a relatively new practice and there is little (if any) documented good practice or guidance available.

- iv. Architectural: systems are **not designed for publishing** open data. Often they will have been developed *ad hoc* for a single operator, making the data difficult to access and extract.
 - v. Public concerns of **privacy and accuracy**: data may need significant amounts of “cleansing” or anonymising before it is published.
 - vi. **Unrealistic expectations or dependency** from the public around the local authority’s capacity to provide consistent, convenient and reliable data all the time; while the local authority will develop and release data, it is in essence putting a data “by-product” to good use – the core function of the local authority remains the management of public assets on behalf of the public, with the common good in mind.
4. As a reasonably new phenomenon, there is very little documented **good practice or recommendations** on how to go about opening up transport data. The lack of a practical framework to guide authorities in making data available has been identified by the Rotterdam Open Data project, a consortium of municipal bodies, universities and local businesses set up to understand the opportunities and implications of opening up Rotterdam’s public data.
 5. A number of concerns have been raised by transport authorities relating to the **cost of opening up data**, which is very pertinent in the current climate of public sector cuts and in view of the fact that most authorities do not have a dedicated budget for their open data activity. This cost does not just relate to building and providing the open data facility but also relates to the ongoing costs of **maintaining open data** (ensuring that authorities have the resources to update/refresh the data once it is published) as well as the support that must be provided to the developer community.
 6. Local authorities implementing open data do so primarily on a free-of-charge basis, ie, no charge is made for accessing the data in most cases. This situation may change in the future; authorities may well start **charging for data** (i) to meet the cost of setting up and operating the open data system; (ii) to ensure improved data reliability and resilience; (iii) in case a policy to commercialise assets is adopted.
 7. The current model for open data presumes free-of-charge data from local authorities, with no subsequent liability for either loss or profit from use of the data. However, the area of **business models** needs to be investigated in order to understand the costs and benefits of opening up data. For local authorities, the benefits need to be considered from both a societal and financial perspective, and so endorsed by the management and/or elected members, should open data supply become a core duty within the organisation. . The province of Noord Brabant is currently developing a business model which seeks to understand the costs and revenues generated by each step in the process of opening up data for multimodal travel information. Findings from a draft evaluation of London’s open data, commissioned by the UK Department for Business Innovation and Skills, showed a return on investment of up to 50-1 in terms of customer time saved.
 8. The technical part of opening up transport data is a small hurdle compared to the institutional aspect. **Institutional and government arrangements are key to progressing open data:**

- i. Central government has a pivotal role to play in putting open data on the national agenda. In the UK for instance, transparency and open data represent the third largest policy plank of the incumbent government. The government has introduced a presumption of publication of public data, meaning that an authority must make a case not to publish data. It has also managed to get a commitment from the deregulated bus and rail industry to publish timetables and real-time information. In the Netherlands, the National ITS Action Plan has been the main instrument for instigating a national policy on open data.
- ii. Central government therefore has a key role to play in creating a framework for open data by means of policy and approach (eg, use of standards).
- iii. A local champion is also very important in bringing about a cultural change within an organisation. London's open data policy, for instance, was championed by the Mayor and enjoyed the support of relevant senior managers. The support of key decision makers has also been instrumental in other cities such as Amsterdam and Rotterdam.

9. Standards

A simplistic approach to open data involves each authority, and each system, simply making available the data that it has stored within it – in whatever form and structure it happens to be. This may be of little use to service providers, who may struggle to understand, integrate and present the data in a coherent form. A preferable solution may be to seek publication in a standardised form. However this will inevitably impose some additional cost and effort on local authorities.

There are a wide range of standards already in existence. It can be a challenge to determine the most suitable standards to use for any particular local authority dataset. For example, an official CEN standard is likely to be comprehensive, but may be excessively complicated and include many different options.

i. Formats

Industry standards and exchange formats are moving much quicker than EU standards. For new data sources (new systems, kit, etc.), it will be necessary to determine what the market will use by way of formats; in this regard, what the market uses could be significantly different from EU-developed CEN type standards (which have a very long lead time). However, the proper procurement of the system could integrate the data export element, reducing the costs associated with providing this data.

For existing data sources (older kit, inventories, manual surveys etc.) there is likely to be significant cost involved in data cleaning and presentation in export formats for the market (which were not required when the data was procured). It would seem reasonable that the data re-formatting costs are borne by, or shared with, the person who wishes to use the data.

Some open data practitioners have suggested that national 'de facto' standards are likely to emerge. In the UK for instance, at least one authority has already adopted the open data formats developed by a large transport authority to enable interoperability and minimal adjustment from app developers.

ii. Data quality

There is always a risk that data gathered for one purpose, for which the quality was sufficient, will be used in open data for a purpose, for which the quality is not sufficient. The quality issue is all the more pronounced for real-time information apps such as passenger information and parking availability since these apps also require accurate geo-referencing data regarding location of the bus/tram stop or car park. Accurate data is not commonplace. Interestingly, there have been cases where the app developer community has helped local authorities improve the quality of this data.

Increasing data quality may therefore be necessary in some cases to avoid disappointing the travelling public and rejection by the app developers. Authorities experiencing this may well consider investing in raising the quality of data.

The data quality risk could also be managed by some sort of labelling/metadata/category system, coupled to a description of the data quality and possibly the original purpose (thereby transferring the risk to the person using the data).

iii. Communications service

In publishing data it is possible to separate out two aspects: the structure of the data messages themselves, and the communications channel through which they are accessed.

Whether or not data format are standardised, it may be possible to standardise the way of accessing or communicating data. This will typically adopt current internet technologies, since these are widely adopted and naturally suited to open access. Current access standards include web services (as a standard front end solution) and xml as a way of describing data structures. In this scenario, a company entering a cloud of public data would find a series of web services, that when queried would return an xml-data file.

iv. Usage conditions

On top of this, many local authorities also include a general statement – typically as a “licence agreement” – which clarifies the basis on which the data is made available. This gives legal protection to the local authority, and may include provisions such as:

- a. The authority does not promise any level of service, and the use of the data is wholly at the user’s risk;
- b. The user may not represent himself as acting as an agent of the authority; or
- c. The authority exonerates itself from any profit or loss or legal challenge arising from use of the data

A standard licence would be helpful for service providers seeking to use data from multiple sources. For example, the conditions established by Transport for London are now in widespread use in the UK.

C. Open data and the EU's ITS Action Plan and Directive

The ITS Action Plan and Directive were adopted at a time when the open data movement was relatively small and not widely known nor understood. Polis believes the open transport data movement could go some way to meet the 'access to data' principles of the ITS Directive. Transport for London, for instance, has released more than 30 data feeds and APIs, which are being used by more than 5000 registered app developers. One app alone has more than one million subscribers. The Madrid bus operator, EMT Madrid, has set up the EMT data platform, which has already generated some 20 apps with more than 350 000 consultations per day. Many more local transport authorities around Europe area are releasing their data and some are taking actions to promote the take up of this data by the developer community, such as Amsterdam and Manchester.

D. Recommendations to the EC

1. The EC should work with a representative number of city-regions and smaller municipalities in a funded programme of applied research, to provide a "Phase 1" set of actions and guidance. Within this arrangement, the real costs and issues associated with open data will be fully exposed, and any subsequent business model will be realistic.
2. The EC should identify the extent to which the growing open data movement is meeting the objective of accelerating the deployment of ITS which is the premise for the ITS Action Plan and Directive.
3. The EC should encourage and facilitate the adoption of a national open data policy by Member States, supported by some start-up targeted investment, to establish an initial framework for open data, coupled to formats and standards.
4. The EC should support the open data movement by facilitating the sharing of knowledge and good practice, and supporting the development of a practical framework to guide transport authorities in opening their data, including aspects such as contractual arrangements, licensing models, business models, realised benefits and challenges, etc.
5. In the context of the ITS Directive:
 - i. The EC should not be over-zealous in the adoption of new regulations in the fast changing area of opening up public sector transport data.
 - ii. The EC should not mandate the use of data standards in the absence of comprehensive evidence of their suitability for open data.
 - iii. The EC should avoid regulations which may benefit the large, multinational information service providers to the detriment of local and regional initiatives.
 - iv. The EC should adopt a focused and balanced approach regarding the transport data to be opened up, enabling data to be opened up gradually and starting with data holding mutual benefits for both the public body and the developer community/market.



EUROPEAN CITIES AND REGIONS NETWORKING
FOR INNOVATIVE TRANSPORT SOLUTIONS



This paper was produced within the Polis Traffic Efficiency & Mobility Working Group with the support of UTMC.

For further information, please contact: Suzanne Hoadley, Polis, shoadley@polisnetwork.eu