

The Open Transport Data handbook

Preamble for contributors - To be removed in final version

What is it

As you may have been able to understand from the title, we are writing an Open Transport Data Handbook: everything a transport agency should know about how to open up their data, how others did it and the possible added values to the company and its customers.

Transport agencies should be defined broadly: we're talking about trains, buses, but also traffic, parking lots, ferries, planes, and so on.

How are we going to achieve this

At the Open Knowledge Festival 2012 we gave a kickstart for the open transport data handbook. We came up with a timeframe and content scheme to collaboratively write the entire handbook. Our soft deadline is the first of November, for both graphic design and content. This means a first version of the text needs to be done the first of October, the graphic design will be done in the middle of

After the contributors section (feel free to add yourself to the list), the table of contents are given. The link of other chapters are added after every chapter.

TODO:

<http://opendatastories.org/2011/06/opening-up-transport-data-creates-economic-and-social-value/> → reuse this text, it's awesome

And this:

<http://blog.programmableweb.com/2012/07/04/73-transportation-apis-yahoo-traffic-irail-and-tfl-cycle-hire/>

Contributors

- Kat Borlongan, @katborlongan
- Daniel Dietrich, @ddie
 - Methodology
 - General advice
- Pieter Colpaert, @pietercolpaert
 - General outline

- Why open transport data
- Miet Claes, @choisissez
 - Graphic Design
- Peter Hicks, @poggs
 - Experiences of OpenTrainTimes
- Samuel Goëta, @samgoeta
 - How does open data change transport organization?
 - MBTA (Boston) open transport data
- Stéphane Guidoin, @hoedic
 - Traffic data
- Tero Piirainen, @teropiirainen
 - 3.2 What data are we talking about?
 - 3.3 Transport companies that are opening up data (Experience in Finland: Helsinki, Tampere, etc)
 - 4. How to open up Transport Data
- Julian Tait, @julianlstar
 - Transparency in charging and why it is so important for multi-modal travel
- Andrew Byrd, abyrd@openplans.org
- Marc De Vries
 - Open Transport Data Manifest
- Nicolas Saunier https://mastodon.social/@aphex_twin/
- *add yourself*

People to ask for a contribution

- People from workshop at OKFestival
- Open Plans
- *add more*

People who should be asked to revise a first version in 3d week of October

- Martin Kaltenböck
- Ton Zijlstra
- Daniel Dietrich
- Rufus Pollock
- Daniel Rudmark
- *add more people*

People who should be sent a free copy the first of November

- All contributors
- SNCF transilien
- Pieter van Ostaeyen, De Lijn, Belgium
- Marleen Verheyen, Infrabel, Belgium
- Peter Van der Perre, ITS Belgium, Belgium
- Mark Farrow, Head of Transparency, Network Rail, UK
- Barry Chesterman, IM Delivery Manager, Network Rail, UK

- *add more people*

Foreword/quotes by Siim Kallas Cabinet?

To be confirmed

http://ec.europa.eu/commission_2010-2014/kallas/index_en.htm

“A very cool quote”

~ Siim Kallas

Table of contents

[Note: Feel free to start editing right away](#)

[Preamble for contributors - To be removed in final version](#)

[What is it](#)

[How are we going to achieve this](#)

[Contributors](#)

[People to ask for a contribution](#)

[People who should be asked to revise a first version in 3d week of October](#)

[People who should be sent a free copy the first of November](#)

[Foreword/quotes by Siim Kallas Cabinet?](#)

[Table of contents](#)

[1. Introduction](#)

[Credits](#)

[2. Why Open Transport Data?](#)

[2.1 Open Data](#)

[2.2 Open Transport Data for organizations](#)

[2.3 Open Transport Data for society](#)

[Sustainable mobility and public transport](#)

[Traffic and parking spaces](#)

[Bike and pedestrians](#)

[3. What is Open Transport Data?](#)

[3.1 Defining “open”](#)

[3.2 What data are we talking about?](#)

[3.3 Transport companies that are opening up their data](#)

[How does open data change transport organizations?](#)

[Network Rail, Great Britain \(WIP\)](#)

[Experience in Finland](#)

[4. How to open up Transport Data](#)

[Legally: Apply an Open License](#)

[What Legal \(IP\) Rights Are There on Transport Data?](#)

[What licenses are already there?](#)

[Technically: Make Data Available](#)

[Methods](#)

[Identifying real-world objects](#)
[Technically: choose standards when possible](#)
[From a community perspective: Make data discoverable](#)
[5 stars of Open Data Engagement](#)
[Examples](#)
[Glossary](#)

1. Introduction

Dear reader,

We have not created a best practices catalogue. We believe every country, government or organisation is different and should use an approach that suits them best. We thought this book can help in the process of implementing an open transport data policy by covering all the aspects and making experts talk about their experience in intermediate cases.

We have not written this book in an ordinary way. It has been conceived and created by a team that is working with (open) transport data day in day out. The team consists of a variety of people from different countries, each with their own background. Some of them will continue working together at the open transport data working group of the Open Knowledge Foundation (OKFN).

OKFN is an international organisation founded in 2004 to support open knowledge all around the world. With the Open Transport Working Group, we want to bring together the goals of transport agencies, policy makers and the reuse community who build upon open transport data. We hope this handbook is the first of a lot of useful resources for open transport policies.

Yours sincerely,
The authors

Credits

- copy contributors from the top + add copyright and licensing information

2. Why Open Transport Data?

2.1 Open Data

"A piece of content or data is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike."
«The Open Definition (opendefinition.org) »

The concept of Open Data has been around for a few decades in different forms and for various

reasons. For governments it was a logical consequence of the freedom of information act, which allows any citizen of a government to request, for instance, collected statistics on demographic growth, climate, education...

In Europe, the term open data was only coined a few years after the Directive 2003/98/EC of the European Parliament and the Council of 17 November 2003 on the re-use of *public sector information* (PSI) (\cite{<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:EN:PDF>}), which sets out the rules and practices for accessing public sector information resources for further exploitation¹. It is the start of open data as standard practise, where not implementing an open data policy in the public sector has to be substantiated.

When, thanks to the freedom of information act and the PSI directive, data would be released on request regularly, a data portal would come in handy, as interested parties requesting data could first look it up on the public data website. A lot of government administrations are now publishing their data on public data portals, such as data.gov, data.gov.uk, opendata.go.ke, data.overheid.nl, and even local governments such as data.wien.gv.at or data.gent.be.

This is more or less where we are today (~ November 2012), but it is not where it is going to stop. The concept of open data implies that data are machine-readable. This means that not only people can have a look on these data portals, but that also machines can automatically discover new datasets. We will thus publish data without password or key protection, we will publish all data using unique URIs² returning raw data in a mark-up language of choice (json, xml, csv, yaml...).

Once the data is accessible by machines, we will get an entire web of data and even things, rather than a web of sites. Instead of querying a pool of web pages, we will be querying worldwide data which will be indexed by third parties, leading to everyone using the latest data which someone published.

2.2 Open Transport Data for organizations

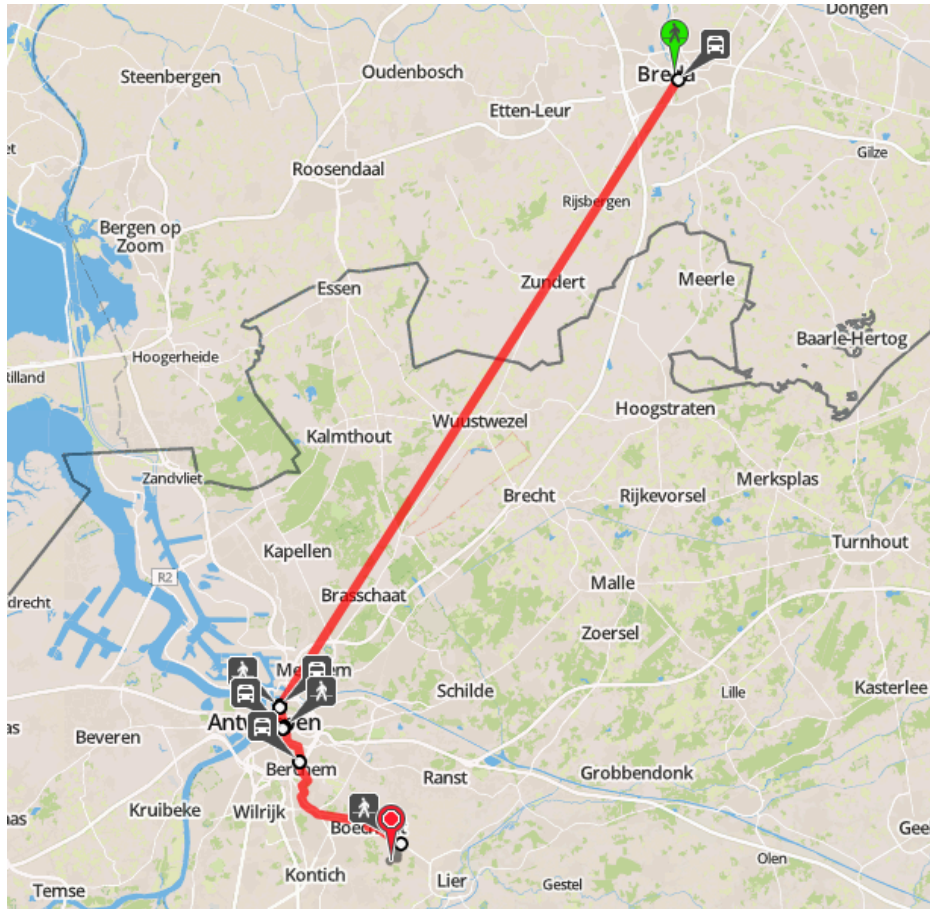
Case 1: an intermodal trip planner

The quest for an intermodal system: a smart mobility challenge
http://ec.europa.eu/transport/its/multimodal-planners/index_en.htm.

The biggest problem is not the technology, but it is the data:

¹Sentence copied from
http://en.wikipedia.org/wiki/Freedom_of_information_laws_by_country#European_Union

² Uniform Resource Identifier: basically an identifier of a dataset using a WWW address



This trip planner feeds of public data, and it runs on open source software. Let's break it down:

- Map Data: open street maps
- Software: Open Trip Planner
- European transport data: open data from The Netherlands, the UK, Paris, Helsinki and data requested under private licenses from De Lijn.
- Open Standards: GTFS, GTFS real-time

//De Lijn - Pieter Colpaert

//This part below needs a sort of introduction

- * Increase ticket sales
- * Better rush hour control
- * Improve the flow of information to passengers
- * Allow developers to fill gaps in the marketplace
- * Use of existing tools (eventually open source) when using some existing standards
- * ...

For traffic data:

- * better cooperation between public institution, for example for the planning of road construction/maintenance (in many places, cities have difficulties to know when a higher jurisdiction like regional agencies or state will do some roadworks)
- * Ability to take into account data from other sources to do internal choice (for example data about road closure/construction could help transit agencies to plan changes of bus trip)

!Note: Not all transport agencies want to increase the number of people on their system.

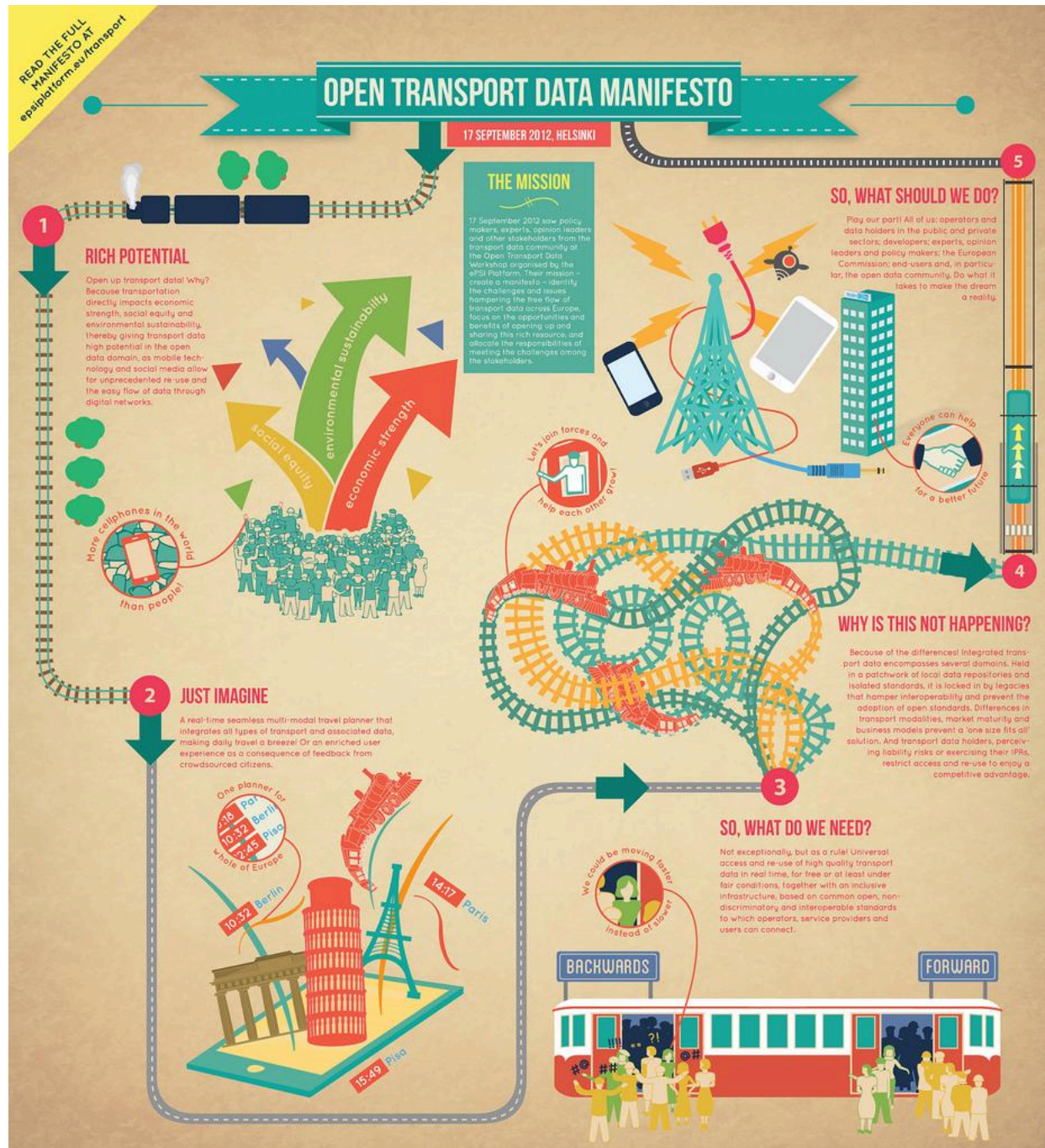
2.3 Open Transport Data for society

Currently, transportation data tend to be split in three categories: transit data, traffic data and active transportation-related data (pedestrian, bicycle, etc). Given that transit is usually operated by public bodies, the use of open data is already wide-spread. This is why existing experiences are dominated by transit experiences. Government agencies have some traffic data, but the most interesting pieces are collected by private entities which makes it more difficult to open data, but it is not less useful.

Case 2: The Open Transport Data Manifest

On 17 September 2012 the ePSI Platform organized an Open Transport data workshop, bringing together an informal group of 30 policy makers, experts, opinion leaders and other stakeholders within the transport data community, from over 15 countries. Sharing a mission to maximize the benefits of opening up this domain, they contributed to the establishment of this 2012 ePSI Open Transport data Manifest, which captures and highlights the potential at hand, the obstacles and the solutions at hand and accordingly allocating the responsibilities to meet these challenges to the appropriate stakeholders.

As such, it seeks to stimulate dialogue between stakeholders by identifying transport data specific issues, that are hampering the free flow of the Open Data across Europe and convince and support (semi) public sector board transport data holders, demonstrating the richness of sources and the impact Open Data may have on their organisations and subsequent value chain.



Sustainable mobility and public transport

Our population is growing. It's a fact that in a certain amount of time we will be with 10 billion people [Rosling image] and this raises a lot of questions. For instance, "how are we going to get all mouths fed?", "is our society going to support this many people?", or "how is everyone going to get from one place to another day in day out?". It's the latter that spikes our interest the most.

The global world population does not act as 1 body. Every single person's actions can be explained from his or her direct situation. [cite MJ O'Laughlin - 2002] When this person has to decide on a transit mode for getting from his current location to another, the best option will not always be the most ecological or the lowest CO2 emitting, but often the most comfortable or the most convenient. The only way to change the way the average person acts is to increase his comfort and increase incentives.

//paragraph 3: information delivered through most convenient channel:

//paragraph 4: For instance a bus company can only innovate as long as there are people on their buses. When certain lines that are stopped when there are suddenly no people on it, it will be even harder to get people on the line if it would be added again in the future: people are not sure anymore it actually exists and miss-information is spread mouth-to-mouth. While less people are on lines, the more people are going to switch to alternatives

Traffic and parking spaces

Although there is a general push toward transit and public transportation to decrease road load and environmental effect, cars and trucks will still be needed for the coming decades and open data could help alleviate a disease faced by almost all the cities around the world: congestion with its related problems: environmental impact, economic losses and social issues (stress, liveability of cities, etc.).

The coverage of traffic data is wide. The obvious issue is the ability to provide real time information to road user to help them plan their trip (or let a trip planner handle this). To get there, we need real time traffic data (load and speed on the roads), ongoing and planned events (accidents, construction, extreme weather conditions). Improving the process of looking for a parking could also be improved by providing open data about the load of parking spots. Various actors are interested to get traffic related data like traffic counts, fatalities and historical data of road load. Finally, most of this data rely on getting the road network of cities and countries to map the data.

Currently, some parts of this data is created by private corporations. For example companies like INRIX, NavTeq, Tomtom, and others are collecting historical and realtime data thanks to crowdsourcing and floating cars data. Public bodies as well as private companies also have some sensors on some road networks as a complement. This private/public duality raises some specific challenges to open data since governments do not own all the data available and private companies may be reluctant to publish the data given its underlying value.

On the other hand, reaching the full potential of traffic data will probably require to have open data, at least for the data owned by governments.

(Also: car sharing and car pooling. Could be implemented by private or public organization. In any case, making data public via files and APIs could make it easier for people to use these options and to quantify the gains (e.g how many people and for what kind of trip they are using car sharing/pooling)
(Also: mention the future of driverless vehicles and connected vehicles (vehicle-to-vehicle and vehicle-to-infrastructure communications))

Bike and pedestrians

//Shouldn't we also cover "active" transports like bike, walking, etc. From what I see cities that want to

publish transpo-related data are also look at those fields (like bike parking, bike sharing schemes, etc.)

In order to make cities more liveable and rely less on cars, more and more local government are pushing their citizen to switch to active transportation, mainly biking. The development and success of bike sharing system is a sign that the demand is here.

Many hurdle could be removed by open data. One of the limitation to use bicycle is the lack of security (or the feeling of lack of security). By opening accident data including bicycle, government could show help citizen to decide which trip to choose. Same thing for road quality.

3. What is Open Transport Data?

3.1 Defining “open”

(source: opendatahandbook.org)

Open data is data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike.

The [full Open Definition](#) gives precise details as to what this means. To summarize the most important:

- Availability and Access: the data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. The data must also be available in a convenient and modifiable form.
- Reuse and Redistribution: the data must be provided under terms that permit reuse and redistribution including the intermixing with other datasets.
- Universal Participation: everyone must be able to use, reuse and redistribute - there should be no discrimination against fields of endeavour or against persons or groups. For example, 'non-commercial' restrictions that would prevent 'commercial' use, or restrictions of use for certain purposes (e.g. only in education), are not allowed.

* data directly accessible over Internet

- not behind log-in
- not behind api-keys

* Raw data

3.2 What data are we talking about?

Data is needed in people and goods transportation whether it is done by cars, buses, trains or other means. Very basic data includes timetables and travelling routes with stops, but there are very many other kind of data that is needed for more advanced transport services.

Open transport data can be categorized as follows:

Planned transit data

- Routes and stops for buses, metro, tram, train, etc.
- Planned route timetables

- Planned stop timetables
- Fares.

Real-time transit data

- Real time vehicle locations
- Real time stop timetables
- Alerts: Delays, Changed routes, Service exceptions, General messages.

Journey Planning

- Journey search
- Stop search
- Journey details
- Journey fare.

Maps

- Raster maps
- Routes and stops
- Points of interest.

Historic data

- Weekly/monthly/yearly reports
- Cultural heritage
 - Many transport organisations have existed for a long time and have big cultural heritage databases including pictures.

Contextual data

- Road network with metadata (speed limitation, addresses, etc.)
- Traffic data (counts and speed by type of vehicle)
- Road closure, constructions, etc.

Data from public transportation has most often been limited to planned routes, stops and timetables, but now more and more agencies are opening interfaces to their Journey planners and real-time data, making it easier to develop third party applications based on those data.

3.3 Transport companies that are opening up their data

//only successful Open Transport Data stories here

- * Experience in New York [OpenPlans]
- * Experience in UK [open train times - Peter?]
- * Experience in Finland [Tero Piirainen, Jari Honkonen, Forum Virium?]
- * Experience in Sweden [Daniel Rudmark?]
- * Experience of SNCF [Kat Borlongan?]
- * Experience of several DOTs in the US for traffic data (PeMS in California, ITS portal for Portland, NGSIM data from FHWA)

How does open data change transport organizations?

To a large extent, releasing data can change the relation of a transport organization with its stakeholders. The open data initiative can also impact the in-house relations between the divisions of the organization as it requires cooperation and involvements of teams. To run smoothly, the organizational side of open transport data initiatives needs to be carefully thought through as it might generate misunderstandings and requires processes of coordination.

Open data initiatives must be clearly explained as they might generate misunderstandings inside transport organizations for several reasons:

- data is considered as a valuable asset for the organization: releasing it for free could be considered as a loss;
- when released, reliability of data can't be completely assured;
- open data initiatives will require involvement of several departments inside the organization (communication, information system, legal department) and requires both coordination and time;
- open data will eventually empower civil society and businesses at the expense of in-house tools: for instance, trip planning can be handled by search engines instead of in-house tools.

These considerations can slow down the data release process, so focus has to be made on explaining the benefits of an open data initiative for the organization such as:

- open innovation: new services can be offered to the public at no cost;
- transparency: a better image of the company;
- objectivity: data can help explaining and understanding a transport organization's work.

Eventually, to make sure, open data initiatives run without troubles, transport organizations need to settle several processes during the process of opening their data:

- host workshop to explain the open data initiative;
- settle a steering committee which gathers different components of the organization;
- run an internal survey to test if employees understand why and how the organization releases the data;
- publish an internal video which explains the process;
- include the name of the data producer so that data can be traced.

Network Rail, Great Britain (WIP)

The rail industry in Great Britain is split between Network Rail, a not-for-profit corporation who operate the infrastructure, and a number of train operating companies (TOCs) who run services. Franchised TOCs are a member of the Association of Train Operating Companies (ATOC), who provide a set of central services to their members - including the National Rail Enquiries (NRE) service.

The NRE website contains real-time train running information through the Live Departure Boards service, as well as service alterations, disruptions and planned engineering work, and is the primary source of information for rail passengers.

There is no API to this data which meets the criteria of the Open Data Handbook. Instead, NRE issue licences - at varying cost and with a standard set of terms and conditions - and an API key to those who people who meet their criteria for access.

A 'push port' is available, which contains a stream of data, but access to this system is restricted and NRE's Code of Practice (<http://www.atoc.org/about-atoc/national-rail-enquiries/code-of-practice>) outlines that a financial indemnity is required in case

The requirement to apply for an API key is a contentious one, to the point that the Office of Rail Regulation, the body responsible for enforcing competition law in the rail sector, opened a consultation (<http://www.rail-reg.gov.uk/server/show/ConWebDoc.11063>) on Real-Time Information in December 2012.

NRE's key concerns are that transparency costs money, the increased load on their systems, and the possibility of an error in a developer's use of data causing an impact on a TOC's revenue and reputation.

History

I was one of the developers interested in using the API when it became publicly known in 2010, but after reading the terms and conditions, I decided not to apply for a licence and instead, to see what I could do to work around the problem and get free access for everyone.

The system which powers the Live Departure Boards service, Darwin, takes data from systems at Network Rail, the Customer Information System at some stations, and allows manual input of delays, cancellations and train amendments.

After securing a regular feed of timetable information from Network Rail, albeit in a rather obscure format, I wrote an Open Source web-based application to import the timetable information into a database and allow it to be queried.

Several months later, I asked Network Rail for access to their real-time information systems through their extranet. As nobody outside the rail industry had asked for access before, they needed to discuss it internally and make a policy decision - however within a few weeks, they responded with a 'yes', much to my amazement.

After jumping the technical hurdle of learning about message queueing, I spent four months work making sense of the data - much of the data is from automatic train reporting systems and needs 'tidying up', but it is useful nonetheless.

Toward the end of 2011, the Department for Transport's 'Transport Transparency Board' announced that Network Rail would be opening up these data feeds to the public in 2012. This was an unexpected, but substantial win, which solved a number of problems and opened the door for further openness within the rail industry.

In early 2012, Network Rail appointed a Head of Transparency, Mark Farrow. Soon after, they released a number of other data sets including PDF versions rail geography data. Summer 2012 saw the release of the datafeeds under a permissive licence in mid-2012 at <http://datafeeds.networkrail.co.uk/>.

Throughout my work, I kept in touch with people who were interested in what I was doing. This is an important part of getting data open - being open about it.

Experience in Finland

Open data is a relatively new phenomenon in the Finnish transportation. However, there is a growing group of businesses and private operators who have started to utilize open data as part of the new services, including several public transport operators like Helsinki Region Transport HRT and City of Tampere Public Transport.

In a relatively small market, it has been challenging to establish good public transport services, and even more challenging to create localized online and mobile services. Open data has been seen as one of the solutions, to provide new online and mobile services with reasonable cost.

Helsinki Regional Transport (HRT) is responsible for the planning and organizing public transport services in the Helsinki region. One of HRT's goals is to increase the modal share of public transport through a system that is efficient and attractive for residents, commuters and visitors.

HRT's route planning application Journey Planner (<http://www.reittiopas.fi/en/>) has served users of public transport already for a decade. From year to year, it is one of the country's most popular online services. In the brand survey 2011 it was the second most popular online service in Finland right after the public library services; just ahead of Google Search.

Still, during the decade of operation, the Journey Planner did not reach all of the potential users. It was targeted to desktop computers, and providing high quality native applications to all the mobile platforms seemed an impossible task. The proposed solution was to open the transport data for developers.

In 2011 HRT opened the transport data and launched an application competition, which inspired the developers to code dozens of applications for public transport passengers. More than 60 mobile applications participated in the competition. Many of the apps are being developed further and now, only 18 months after the competition, the HRT public transport information can be accessed with number of convenient mobile apps, available for all the mobile platforms. Not all of the applications are conventional route planners; there are also applications for sightseeing, gaming and urban adventures as well as traveling assistants for the visually impaired.

According to surveys, opening the data has improved the service to users of public transport. The results are excellent especially if we take into account, that the costs of development and maintenance of the open interfaces have been only a small fraction in the costs of the information system.

Encouraged by the success in the Helsinki region, also other cities in Finland are starting to take similar approaches, and opening up their transport data sources. The smaller cities like Tampere, Oulu and Turku are closely following what interfaces and standards are being deployed in the Helsinki region. Having the same interfaces and standards available in other cities, will make the costs lower for the public authorities, as well as life easier for the application developers.

Transit and traffic in the San Francisco Bay Area

[Waiting for comments from MTC and BART]

Two leading organizations are providing an interesting insight about open data in the transportation world. Both Bay Area Rapid Transit and the 511.org initiative started very early to share open their data, usually by opening APIs that were initially designed for internal use. Given that both are located in the San Francisco bay area, which is considered as the epicenter of the startup world, it provides an interesting comparison. California is also plagued by many transportation issues. As a consequence, nobody should be surprised to see a vast ecosystem built on any interesting dataset released by transportation agencies.

BART operates a commuter train network and thus focus on transit data. 511.org provides a wide variety of data (from transit schedule to rideshare system) in a large region but for the needs of the comparison, the focus will be on their traffic data API (network, events and speed profiles).

Although both traffic and transit APIs are made available by each organization for a long time, it is striking to see that BART is the center of a large ecosystem (about 80 applications) while 511.org traffic data is integrated in a handful of applications. Although a comparison is not easy, here are some possible criteria affecting adoption of data by an ecosystem of web and mobile applications.

Standardization: One of the main difference between transit and traffic is the level of standardization dedicated to open data. Several powerful open data formats exist for transit like General Transit Feed Specification (GTFS) and SIRI. Both formats have been in place for a while and are supported by a wide range of commercial and open source software. BART provides standard data feed as well as custom APIs. Consequently, any tool that is able to support GTFS can also retrieve data from BART. On the other hand, the standards for traffic are more designed for center to center communication and transportation experts and less oriented toward open data. As a consequence, there is still no standard commonly accepted to publish open traffic data.

Business complexity: The complexity of the traffic standard is explained by the complexity of the business. While transit agencies have a clearly and segregated network where they have a fair amount of control on their buses and train, traffic is much more complicated: the main roads are managed by states and regional entities while the local network is managed by cities. The possible cases are more complex to model which results in more complex standards.

Merging data: Because of the complex traffic environment, there is a need to merge several sources: on one side there are several sources of "official" data which could come from the state, the metropolitan area, the 9 counties and 80 cities. On the other hand, since traffic organization cannot have see what is occurring everywhere, a large part of the data is coming from the citizens, either officially or unofficially via tools like Waze. Whether one is on a side or another, it appears complex to find and mesh all the data (which is not always open).

Technical simplicity: In order to be adopted by developers and third parties, open data have to be published in a simple and open manner. For APIs, the use of well-known designs like REST make it very easy for developers to start investigating the data and decide what to do with it. For example, BART made its real-time API based on REST which makes it easy to try, even with a simple browser. The 511.org solution has been designed before the goldrush toward REST and was also thought to meet internal needs. The choice of the JMS messaging system fit the internal needs but fits less with the habits of the open data developers.

Licensing: The license and terms of use supporting any open data initiative is critical. The process to acknowledge the terms of use as well as the constraints in terms of reuse, integration with other data and commercial use can have a serious effect on adoption. BART provides some simple terms of use and provides a common API-key to access their service. On the other hand, MTC asks for paper-signed agreement to provide connection credential. From the point of view of an organization or a developer who wants to develop a tool that gets data from several sources, managing complex license agreements can be a major limitation.

Although it is difficult to weigh the various elements an element is clear: in the quest of usable open data some limitations are coming from the domain (lack of open data standardization and business complexity) while some others can be improved by an individual organization (technical simplicity and licensing).

511.org is currently running a project to maintain its leadership on the traffic data side by developing more simple APIs, improving the data license and by leading a standardization process. Hopefully this will help traffic data to become as accessible as transit data.

A manager from BART underlined how powerful open data is. Since most of the tools they built were needed for internal use, there was almost no dedicated cost to open their data. Open data provided an incredible ecosystem of tools that serve the users.

4. How to open up Transport Data

Legally: Apply an Open License

What Legal (IP) Rights Are There on Transport Data?

// Matthias Dobbelaere ([MyLex](#))

Europe:

- * ITS directive
- * PSI directive
- * European Database laws

//End with, a license is needed on these data and is needed for reason XYZ

What licenses are already there?

- * Creative Commons 4.0

Quote from

<http://blog.okfn.org/2012/08/15/cc-license-version-4-0-helping-meet-the-needs-of-open-data-publishers-and-users/>

- * oDBL
- * UK open gov license

Technically: Make Data Available

Methods

Rephrase 3.2 with now the best methods to make the data available. How does

- * API
- * Downloadable data dumps
- * push notification (not frequently used yet)

Identifying real-world objects

[Ask input of datalift people]

Technically: choose standards when possible

The experience in transit data shows the various advantages of standardisation. Transportation agencies are publishing their schedules and other information in standard formats, bringing dozens of application that help transit user to plan their trip and choose the best option for them.

For open transit data, there are currently two open standards to choose from:

- GTFS is a standard developed by Google. The specification defines a common format for public transportation schedules and associated geographic information. GTFS data feeds are already available from more than 100 transit agencies. GTFS-RT is the standard for exchanging real-time information.
- SIRI is a public transit standard managed by a CEN standardization. It is widely used internationally, although not as much in the US as in Europe.

Thanks to standardization, the tools and applications designed for one city/agency can be easily applied somewhere else. But it is not only for users' apps. Open source tools like OpenTripPlanner (multimodal trip planning) or OneBusAway (realtime bus/train location and notification) can be implemented by transit agencies to provide better services to their client at very low cost.

However, in many fields, no open data standard exists. For example for the traffic data, private companies exchange data in non-public, custom formats and some of the existing standard for intelligent transportation systems (ITS) may be too complicated to fit in an open data policy. But as more and more public bodies are willing to publish traffic data which will inevitably lead to de facto or hopefully official standards that can be reused by other data publishers.

From a community perspective: Make data discoverable

5 stars of Open Data Engagement

//intro about the stars and their origin

//Use the 5 stars of open data engagement as titles
//Ask Tim Davies to write this? *emailed*
//What do they mean, and how can they be applied on transport data
Things we should certainly talk about:
* our CKAN instance with CI (Stefan Wehrmeyer + ITS Belgium?)
* Crowdsourcing data
* ...

Examples

//how did others make their data discoverable in the past?
// SNCF → ROMAIN
// OpenPlans
// Trafiklab

Glossary