Open your transport data!

A simple guide for municipalities, city councils and departments towards open data

Implemented by

giz Deutsche Gesellschaft
für Internationale
Zusammenarbeit (GIZ) GmbH

Content

Why Open Your Cities Transport Data?
Benefits and Opportunities

As public transport users across the world become more digitally savvy and connected, they demand intelligent solutions. One of the ways this need can be met is by providing dynamic real-time travel service status information via real-time travel information apps. In order to develop such applications with a high quality standard, software developers require access to specific data sets, such as arrival and departure times, service disruption or traffic structures.

Governments, municipalities and organisations generate data in various forms – documents, records, transcripts, etc. Originally, the created data sets have not been made available to the broader public, mainly due to data protection concerns. However, opening data to the public can have several positive effects. It drives internal efficiency, sparks community engagement and fuels civic technology development. Opening up data can also inspire interoperability between various actors of different levels: inter- and intra-governmental, regional, national or global.

The provision of Open Data has delivered benefits to customers and users by improving their journey experience, and to transport organisations by improving the efficiency of operations and providing a statement of openness and transparency (see examples).

The Helsinki Regional Transport Authority (HSL) Navigator[1]

In partnership with developers and fellow organisations, HSL has developed an intelligent route planner to make public transport easier to navigate – based on Open Data. The planner offers nearest public transport (and cycling) options according to the user's current location. It also integrates real time information, is compatible with most smartphones and desktops and works offline.

The universal architecture of the planner and its Open Data policy makes it internationally applicable. Tests in Tampere, Manchester, New York and Berlin have shown that deployment in a new city is possible within only one week, as a result of the open source code and user-generated content approach with no license fee.

1   http://www.uitp.org/open-source-travel-planning
2   http://www.uitp.org/opening-transport-data-london

Opening up Transport Data in London – Experiences of Transport for London (TfL)[2]

Over 200 travel apps have been created for London transport users since 2007, when opening up transport data began. These apps, all based on TfL's free and Open Transport Data, cover all modes of transport and reach millions of customers. Travel information apps and web-based services developed using Open Data help customers to make better-informed journey choices that save time and optimise the utilisation of the network, estimated at an economic value up to £58m per year.

How to Open Your City's Transport Data

5 Steps towards Open Data[3]

1 Identify Responsible Staff

- Select a responsible department/section/staff to accompany the opening of the data

- Ask for help and expertise if needed. Assistance or capacity development can be provided by several stakeholders (scientific institutions and research foundations, peer-to-peer learning with other cities, consultancies)

2 Choose Specific Dataset(s)

- Only release data which can be classified as "Open Data" - Open Data does not contain any personal information or traceable personal data

- Keep it simple. As a first step, concentrate on the datasets which are already digitalised and available

- Keep innovation in mind. While selecting relevant datasets, look for data that can be useful for developers. Ask the community about their needs.

3 Apply an Open License

- Determine what property rights exist in the data
- Apply an open license that suits your datasets
- Ask yourself :

I Who are your future users?
Ask the community and offer potential datasets to gain feedback and give you an understanding of the demand, and run consultation events.

II What will the data be used for?
Allow data to be used for commercial or non-commercial purposes? To be printed, copied, altered, presented, processed and transmitted to third parties?

III Involvement from third parties?
Create avenues for users from the private and community sector to add their data and facilitate data improvement by allowing derivatives to be catalogued.

IV Access
Ensure that access is equitable for everyone.

---

3   Steps 2-4 are based on Open Knowledge Foundation Handbook
    http://opendatahandbook.org/guide/en/how-to-open-up-data/

## 4 Make the Data Available – Technically and Legally

- Data should be available preferably as a free download of complete datasets.

- Open your data in a useful and machine-readable format. PDFs or other closed formats are not accessible for many developers. Try to create as few hurdles as possible.

- Create an online catalogue for government's data. Free and open source software is also available and has already been widely adopted by governments.

- You may also wish to consider alternative ways of making it available such as via an API (Application Programming Interface)

- Align your datasets/API to already existing ones to make it easier for developers to make use of.

5 Make it Discoverable

- Make your data available online

- Link your datasets to existing open databases

- Promote your open datasets via mailinglists, organisations, social networking groups, hackathons. etc.

6. Promote Open Data in your network!
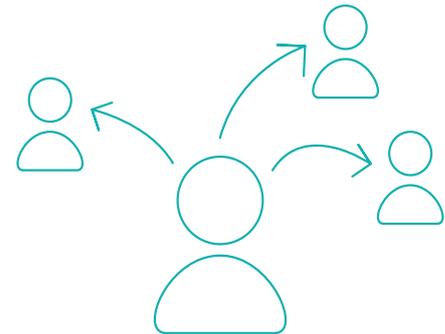
Ruter Labs: A pioneering Open Data Approach for South-Eastern Norway[4]

Ruter is the public transport authority for the Greater Oslo area, where 2.5 million inhabitants benefit from their services (about half of Norway's population). Ruter provides approximately 80% of the country's public transport services.

Prior to 2008, data on Ruter's services was not publicly available, apart from an API for real-time departure information on a public server. Media interest increased in the context of a test project for early smartphones in association with Opera Software. As a result to the test project, the real-time API was made available to developers working on a non-profit basis.

From 2009/2010 onwards, a new technological platform with a common API for real-time information, travel planning and deviation information served all of Ruter's information channels such as the website (http://ruter.no) and travel app RuterReise.

A 2010 initiative from the Agency for Public Management and eGovernment (DIFI) further encouraged all publicly owned companies to release Open Data to everyone and a common Open Government Data license (NLOD) was introduced. As a result, Ruter's Open Data became available through the DIFI's portal, its own Labs site and Facebook community page in 2011.

4   http://www.uitp.org/ruter-labs-pioneering-open-data-approach-south-eastern-norway

Useful Tips

Throughout the process of opening data it is important engage with actual and potential users often and early. Most of the data will not reach the final user directly, but rather via "mediaries" – people who transform and present the data in different forms, such as maps. Such mediaries should be included in the process of user engagement. Additionally, the concept of Open Data naturally raises a number of concerns and questions, including data use, licensing, third party involvement and privacy. Therefore, it is important to be prepared to address common fears, questions and misunderstandings.

A short introduction into APIs

API stands for application programming interface (API) and is a set of commands, functions, protocols, and objects that programmers use to interact with an external system or to create software and applications.[5]   API can ease the programming of specific components by integrating new features into existing applications (plug-in API) and can help to integrate and enhance the functionalities of applications through sharing data.

A popular example for the use of API is the use of Windows API in regards to Microsoft Office. The Windows API enables external firms to develop software for Microsoft Office. Using the API thus makes it possible to develop applications that run successfully on all versions of Windows while taking advantage of the individual features that are unique to each version of Microsoft Windows.[6]

This way, publishing APIs has encouraged sharing data and content between user communities and applications. When content is created in one place it can be dynamically posted and updated in several locations on the web. Examples of that are:

• Sharing photos from Flickr to social networks like Facebook
• Embedding content like YouTube videos
• Dynamic posts – sharing comments made on Twitter with a Facebook account
• Adding features like a Follow Button to blogs or websites

There are two main ways of releasing APIs. Information on APIs can be either protected from the general public so that only licensed developers can access data. In that way the data provider has quality control over the use of data and license revenue streams - making video game console APIs available only to licensed video game developers is an example. The second option is to make APIs freely available. This is the case for Microsoft Windows, as well as Apple's APIs Carbon and Cocoa.[7]
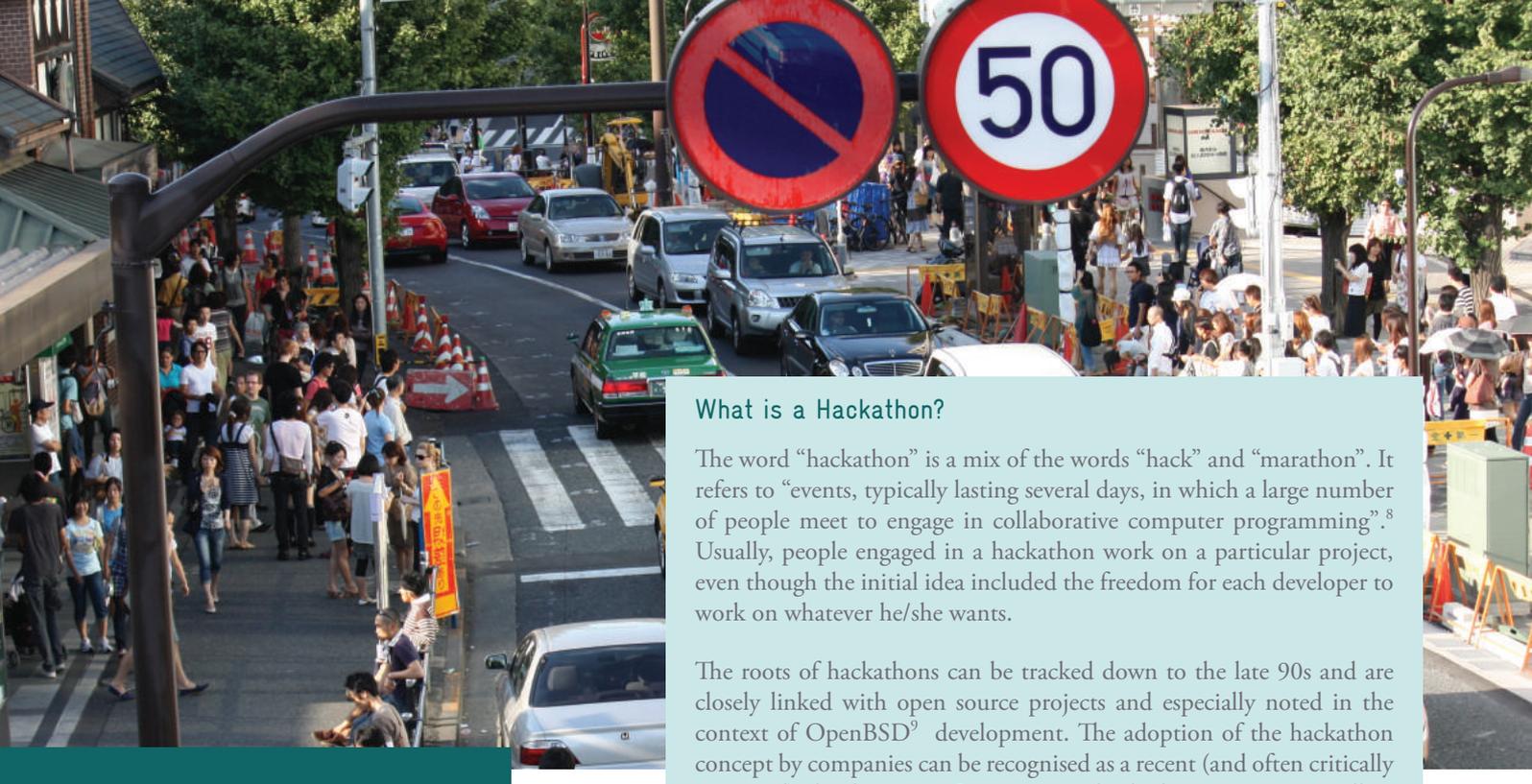
5   http://techterms.com/definition/api
6   https://msdn.microsoft.com/en-us/library/windows/desktop/ff818516(v=vs.85).aspx
7   Tech Target: Essential guide to API management and application integration:
    http://searchsoa.techtarget.com/essentialguide/Essential-guide-to-API-management-and-application-integration

What is a Hackathon?

The word "hackathon" is a mix of the words "hack" and "marathon". It refers to "events, typically lasting several days, in which a large number of people meet to engage in collaborative computer programming".[8] Usually, people engaged in a hackathon work on a particular project, even though the initial idea included the freedom for each developer to work on whatever he/she wants.

The roots of hackathons can be tracked down to the late 90s and are closely linked with open source projects and especially noted in the context of OpenBSD[9] development. The adoption of the hackathon concept by companies can be recognised as a recent (and often critically perceived) development. Some view a hackathon as a way to allow development teams to work on chosen projects.[10]

8   http://www.oxforddictionaries.com/de/definition/englisch/hackathon
9   OpenBSD is a computer operating system descended from Berkely Software Distribution (BSD), a research unit derivate developed at the University of California, Berkeley. The project
      is foremost known for its developers' insistence on open-source code, open documentation, code correctness and security.
10   https://www.techopedia.com/definition/23193/hackathon

Open Data Licenses
Why licensing matters

Open data is vital to fostering innovation and knowledge exchange. A clear definition of Open Data is necessary in providing a standard to ensure the interoperability of different datasets and, in this regard, licensing is key. Open data licenses give users a clear framework of how they are allowed to use the data provided by governments, whether they are allowed to modify it and/or share those changes. Together with a definition of open data, licensing thus ensures the highest level of clarity and simplicity.[11]

Since the concept of open data sets and related licenses may seem rather abstract, it helps to take a look at a more popular form of knowledge sharing licensing in the digital age. When it comes to digital pictures, videos and other forms of creative work, creative commons licenses[12] reached tremendous popularity in regards to open knowledge sharing and a common understanding of its usage.

Creative commons works with flickr, youtube, vimeo, wikipedia and many other partners as creative commons platforms, and offer security and transparency for sharing and using data through their licenses. Concerning entire data sets, the most-used international licensing system is called Open Data Commons. Though different criteria concerning attribution etc. applies, all ODC licences imply that the data sets can be used for non-commercial as well as for commercial purposes.

11 http://opendatacommons.org/faq/
12 https://creativecommons.org/about/

Public Domain Dedication License
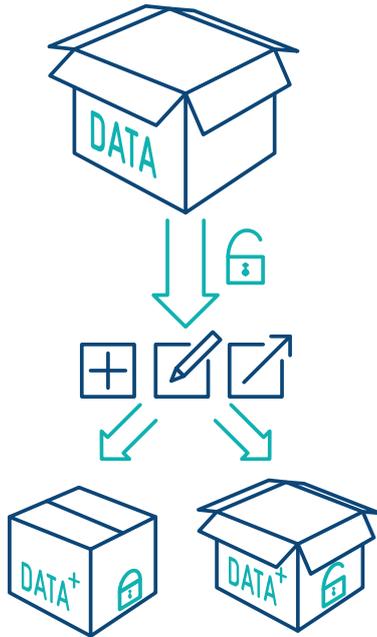
Open Data Commons Public Domain Dedication and License[13]

With a PDDL the data/database gets placed in the public domain, thus waiving all rights. You are free to create, adapt and share from the given data without any restrictions.
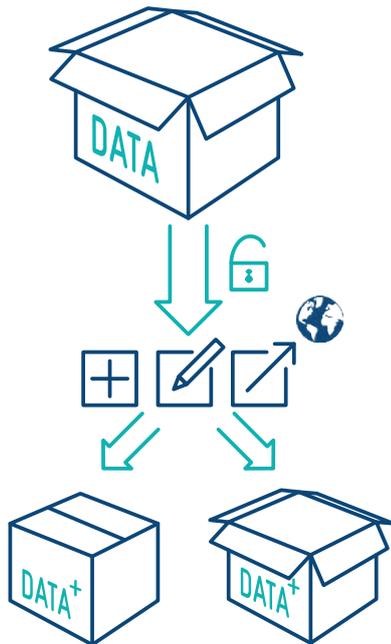
Open Data Commons - Open Database License
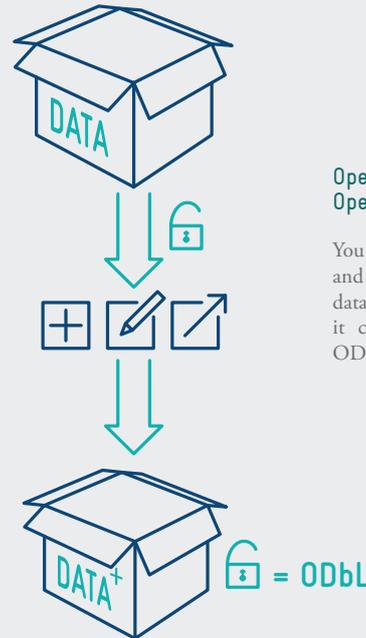
Open Data Commons Open Database License

You are free to create, adapt and share from the given data as long as you attribute it correctly, share it under ODbL, keep it open.

= ODbL

13 http://opendatacommons.org/licenses/

Open Data Commons Attribution License

Open Data Commons Attribution License

You are free to create, adapt and share from the given data as long as you attribute it correctly.

Create

Adapt

Share

Open Data
Package

Closed Data
Package

German Data Licenses – Approach

The German licensing system Data License Germany addresses the question of commerciality purposes differently and distinguishes between commercial or non-commercial for the use of a certain data set and license. However, non-commercial makes the data sets useless for start-ups and can hinders effective developments.

Data License Germany – Zero[14]

Any use is permitted without restrictions or conditions. For commercial or non-commercial use you are free:

- To Adapt: To print, present, alter, and process
- To Share: To copy and transmit to third parties
- To Merge: With own data and the data of others to form new and independent datasets
- To Create: To integrate in internal and external business processes, products and applications in public and non-public electronic networks

14 https://www.govdata.de/dl-de/zero-2-0
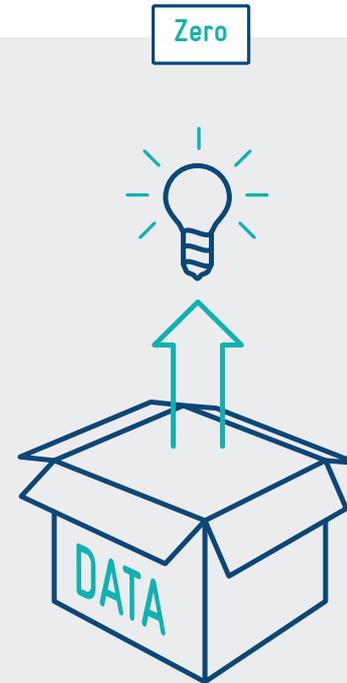
Attribution and Non-Commercial

Data License Germany – Attribution – Non-Commercial[15]

Similar to the Open Data Commons Attribution License (ODC-By) in the sense that you are free

- To change and edit, and
- To add new designs or other amendments

As long as those changes are marked in the source note, or the source note is deleted. However, there is a clear differentiation, as any use for commercial purposes is not permitted. This way, contrastingly to the ODC-By or the Zero data license, this non-commercial license makes the data useless for developers and start-ups, prohibiting any cooperative formation between developers and governments for economic innovation.

---

15 https://www.govdata.de/dl-de/by-nc-1-0

Bibliography

OKF Handbook

http://searchsoa.techtarget.com/essentialguide/Essential-guide-to-API-management-and-application-integration

http://www.oxforddictionaries.com

https://creativecommons.org

http://opendatacommons.org

https://www.techopedia.com

https://www.govdata.de/dl-de/by-nc-1-0

http://www.uitp.org/open-data-reaching-out-our-customers

http://www.uitp.org/action-points-benefits-open-data