Open data - improving border crossing efficiency

Erno Vanhala Uolevi Nikula
{erno.vanhala uolevi.nikula}@lut.fi
School of Industrial Engineering and Management
Lappeenranta University of Technology
P.O. Box 20
FI-53851 Lappeenranta.

Keywords: open data, border crossing, service, business

Open data has been defined in the following way "a piece of data or content 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" [1]. This definition gives the idea how we should handle data state governments are collecting around the world. During the years governments have collected and produced data on maps, car registers, border crossing statistics and much more. Yet, services providing this data are only few – although the number is increasing as governments are releasing their data sets [2].

By giving the data freely available government solves two problems: a) it is not necessary to build services by themselves and b) it gives private companies data they can build their new services upon. This reduces the role of government and lets the markets decide what services are useful and what are not needed.

The process of opening data can have problems as it includes both licensing and technical issues. All the data must be given away freely without any restrictions, so it is not possible to release copyrighted material without contracts. Also the technical side has potential problems as the data should be in an easily handled format. By the definition given in [3] the data released as open can be categorized in different levels. The first level includes only the fact that data is available under open license. It can be in any format (e.g. table full of data given as JPEG-image). The second level improves this as the data is then in machine readable structured format (e.g. excel sheet). The third level is achieved when the format of data is open (e.g. CSV or HTML) instead of a proprietary one. On the fourth level the data is available through URI, so people can access it better. The fifth level links the available data to other data units to provide more value to user. With linked data developer has access to specially encoded data elements to link the dataset to another [4].

So far the Finnish government has released, for example, national maps as open data and the opening movement continues The border between Finland and Russia has a lot of traffic as millions of Russians come to Finland every year and hundreds of thousands of Finns visit Russia. The private border crossing mostly happens in few border stations located in South East Finland. It has been in the news that some times truck queues are tens of kilometers long and travelers need to spend hours on the passport and visa control queue. This has lead to services (e.g. [5]) which show information about traffic issues in the border.

Problems of these services include the fact that the data they show is not clearly licensed as open and it is not machine readable. In the following we present few improvement ideas.

Forecasting queueing time based on statistical and real-time data

The existing services provide statistical data on how long queues exist on a specific day of month. If this data is extended to show queue lengths on a specific day of year and is provided as machine readable data, 3rd party developers would be able to build statistical models to forecast queue lengths and queueing time. If this is then joined with a real-time data the application would be able to give even more accurate prediction on how smooth is the border control at the specific time. This prediction data can then be utilized in trip planning advisor application to provide better estimations for trip time schedules for different time frames.

Opening hours of markets in nearby cities

The data from previous application could also be used in an application to show opening hours of shops and markets in nearby cities. Basically if the border queue is short enough one can get to market before its closing time instead of spending his or her time on the passport control. Besides traveler this benefits also markets as they could stay open longer in those days the customer queue and flow is expected to be longer.

These kind of models express the use of open data to generate cost savings and simultaneously enable innovative new businesses. For the state government it is less expensive just to open the (existing) data than also develop the services utilizing the data. Providing the data in a machine readable format new business can emerge round the data. This encourages to use the data to build new business thus fueling the economy. Figure 1 presents differences between traditional government lead services versus services build by private companies upon the data opened by the government.
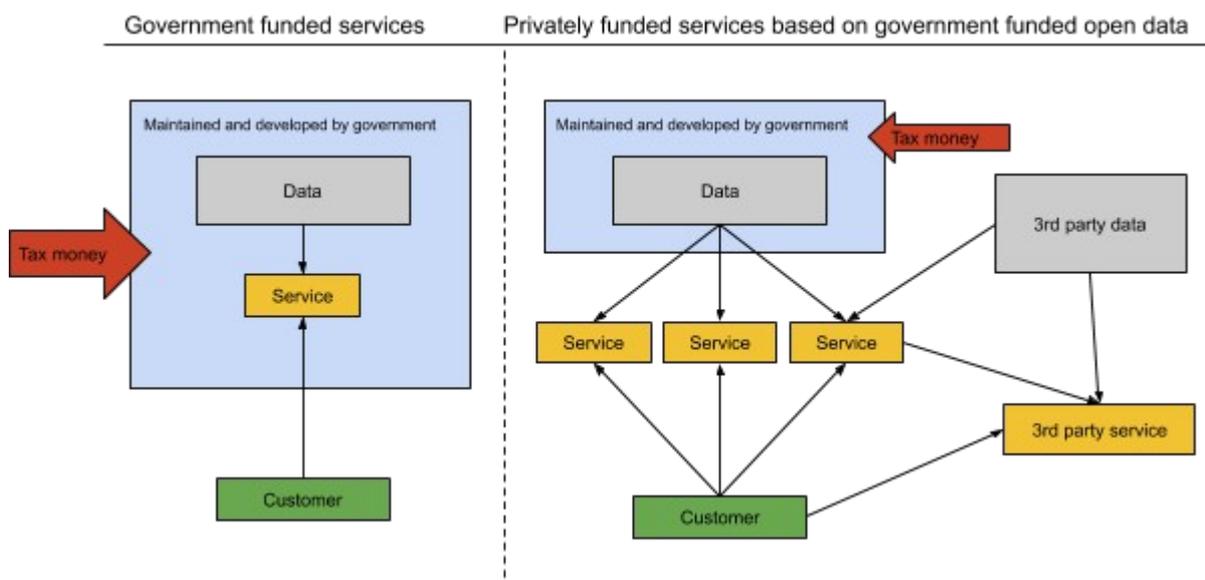

Figure 1: Government and privately funded services.

Figure 1 illustrates the difference between government and privately funded services where the first one is built "in-house" and the second one just gives the data for the software service industry to utilize. The latter gives a possibility to several existing services to compete against each other and thus give customers a choice. These services can use other 3rd party data services and also create

completely new services to be built upon the services using the original data.

As a conclusion we argue that governments should focus on opening datasets instead of developing services by themselves. This reduces the costs and can improve the service quality as individual privately funded companies can build services. For an end-used this means better user experience. Tax money can also be targeted to other areas. In a best case scenario this is a win-win situation.

References

[1] "The Open Definition," *The Open Definition - Open Definition - Defining Open in Open Data, Open Content and Open Knowledge*. [Online]. Available: http://opendefinition.org/. [Accessed: 14-Oct-2014].

[2] L. N. Mutuku and J. Colaco, "Increasing Kenyan open data consumption: a design thinking approach," in *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance*, NY, USA, 2012, p. 18.

[3] "5 star Open Data," *5 star Open Data*. [Online]. Available: http://5stardata.info/. [Accessed: 14-Oct-2014].

[4] E. Kalampokis, E. Tambouris, and K. Tarabanis, "On publishing linked open government data," in *Proceedings of the 17th Panhellenic Conference on Informatics*, Greece, 2013, p. 25.

[5] "Rajaliikenne.fi," *Etusivu | Rajaliikenne.fi*. [Online]. Available: http://www.rajaliikenne.fi/fin/etusivu. [Accessed: 14-Oct-2014].