Kaarel Kõrvemaa

TIMETABLE COMPARISON SEARCH USING OPEN DATA

Degree Programme in Innovative Business Services

2015

TIMETABLE COMPARISON SEARCH USING OPEN DATA

Kõrvemaa, Kaarel
Satakunta University of Applied Sciences
Degree Programme in Innovative Business Services
March 2015
Supervisor: Koivula, Reijo
Number of Pages: 41
Appendices:1

_____

The aim of this thesis is to explain how data can be used to create new solutions. Majority of people don't understand how open government works and what kind of information is available. With this thesis the understanding of open data and solutions that are created with it, became more understandable. The aim is to explain open data with real business cases, which describe and help people everyday.

There are thousands of open interfaces what allow us to connect, search and visualize information that government is collecting everyday. By using this meta data for developing new services has became a phenomena. Practical examples have been used to show what has been done. There are also different options that open data can be used, like research, education, automation and applications.

Solutions called Matkapaiva.fi, which is using Finnish transport agency and Finnish meteorology institute open data application programming interface, to provide a timetable comparison application is a new and innovative mashup. The idea comes from the need to search available transportation types to get from point A to point b. The weather information, which visualizes the temperature in the departure city and the destination, is a combination of different sources that can be used.

CONTENTS

1 INTRODUCTION

This thesis will analyze public administration open data. The theoretical part comes from multidimensional open data process that describes a diverse range of activities that are performed in the open data process (Innovation with open data: Essential elements open data ecosystems, 2014, p.25).

The research question how to combine different open data's into one service will be answered by a solution that implements transportation data and weather data in a useful way. Visualizing data is also an important part of our daily life and creating innovative ways to share it, is public sectors challenge. Therefore the goal of this paper is to explain what open data is, where can it be found and used and also a practical solution, with analysis the users. This thesis combines different articles, case studies and books about how open data works and what kind of services there are.

1.1 Vision

Starting from the beginning it would be wise to start the analyses from libraries. Older days open data meant open information and knowledge. Understanding that books, articles and research were made in a written form, libraries made sharing information available. As of technology and sharing information has become easier, libraries can share their information in data form. This has revolutionized how we consume information. Not only that we can read everything from the web, but also search, reuse and commercialize data. The Open definition sets out principles that define "openness" in relation to data and content (Open Definition, 2005)

In 2007 a group of people in Sebastopol, California had a vision about how democracy could work better, if government data would be open. They had eight principles what data should be and look like: it should be complete, primary, timely, accessible,

machine-processable, nondiscriminatory, nonproprietary and license free. Complete means mostly that the data is open, primary means that data source is clear and not modified, timely comes that is up to date, machine processable means that it's not on a paper, but processable, nondiscriminatory openness that anyone can access data, nonproprietary means that data is not controlled and license free, mean that no copyright, patent, trademark or trade secret is regulating the use or distribution.

Opening government data started in the United States and Britain and has now spread to most industrialized countries. Open data is a philosophy in which the collaboration between the public sector, citizens, web developers and other users is nurtured and made more efficient to produce greater common good (Helsinki Region Infoshare, 2015).

This research is made because of the high interest and growing open data government phenomena around the world. The core is that eventually OGD will increase new business possibilities and social projects around the world. For example European Commission launched an Open data strategy in December 2011, which from their point of view will bring about $53 billion each year to the European economy.

This will only happen when people and private companies will use the possibility to make something out of it. There are two ways to create value in the open data world, one is worth of the service and another is social benefit. OGD refers mostly to the social benefit, which openness and transparency brings for the government.

1.2 Open data

Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and share a like (Open data hand book, 2012).

Open data can be use in cultural, science, finance, statistic, weather, environment and transport. Cultural matter it can be stored in museums, galleries, libraries or other publicly funded and regulated places. Science data that is produced by institutions is

mostly research, for example universities. Finance data produced by governments, central banks and companies, both public and private businesses. Statistic data is mostly produced by statistical offices, most of the countries in the world have their own statistics offices that collect and maintain data files on society. They have the obligation to share that publically.

Weather data is used to forecast the weather. Environment data consists of things that have an impact on nature. Transportation data consists of timetables, routes, road maps and public transportation information. This research will take a look at what is available in Finland in over all and goes deeper into transportation and weather data, which is applied in practice.

1.3 Open Government Data (OGD)

Before explaining the open government data, it would be wise to talk about open government. Last chapter was explaining what open data is, now we will look in to open government and what's the difference between them. Open government gives easy access to government information, the core is not just to get some information but also to get the raw data, how are decisions made. Raw data means meta data. Having meta data from the published information gives more transparency and openness.

Picture 1 shows that the open data and government open data are all connected, so data can't exist, without open government. From the taxpayers point of view, we need to know how our tax money is used, so we could get as much out of it as possible.
.
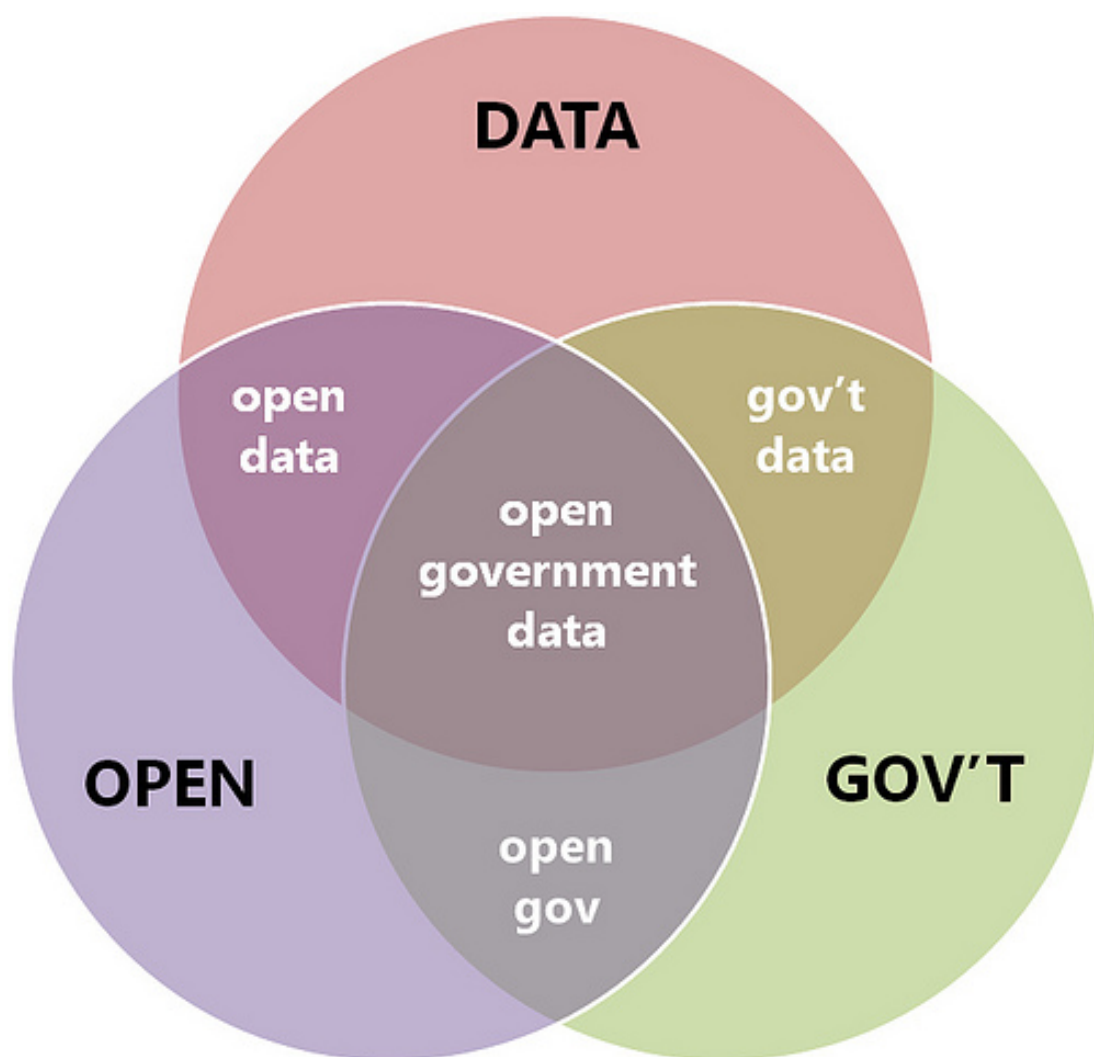
Picture 1. How open data is visualized in open government view

OGD comes from the policies that are made around the world as of the demand for openness in government decision making is increasing rapidly. The core is that the OGD needs to be machine-processable to attract third parties to improve the information flow. Information is a right so providing access is mandatory for governments.

1.4 Open data license

Do use open data it is needed to sign a formal agreement. With the agreement the holder agrees to the terms and conditions. It will give the user incorporeal rights to the data. The license agreement does not create a cooperation or business relationship between the user and the provider of the data. This means that it's free of charge and you will have the rights to the data. Rights mean that holder can copy, distribute, publish, modify, utilize commercially or non-commercially, combine with other products and use as part of an application or service.

With the usage comes duties and responsibilities. When using open data you have to mention what kind of open data are you using and share a link to the main source. When it comes to open data provider, they have to make sure that data is up to date. They also mention that they are not responsible for the correction of the data.

All databases have their own license, license gives a specific data access. Data is stored in most of the cases separately from the main data source. This is for preventing the systems to be safe.

1.5 API

Many articles reference APIs as a revolutional transformation in the communication between the web, not that they didn't exist before, but to what we can do with them now. In the open data world they are the most important ones. It all started from the internet business boom, which started year 2000. It all comes down to your customers, who went there and if you weren't providing service there you quickly lost your customers.

Now it's more challenging to keep up with customers. Old days if you had an website with the information about your business you were okay, but now when users behavior has changed, new devices, like tablets, smartphones and other gadgets, connecting with your customers, has become more complicated. This means that decisions are made all the time, wherever they are.

From the search point of view all the information has to be available, every single word that describes what you do. The data and the information are called digital assets. Digital assets are as important and in some cases more valuable than money itself.

Application programming interface is a level where computer system exchange information. For example, you are about to purchase something from the web. It happens sometimes when you start to check out from the purchases site it will inform that the product you want isn't available. This means that it's out of the stores stock, but it can also notify that it will be available in X days. System uses 2 APIs one is their one and the other one is the producers API.

Picture 2 explains how different stakeholders use APIs, from the user point of view it stays as a one way traffic road, but when it comes to purchasing and exchange information, it can be a two way road, like the web purchase example. OGD systems share information the same way like private ones.
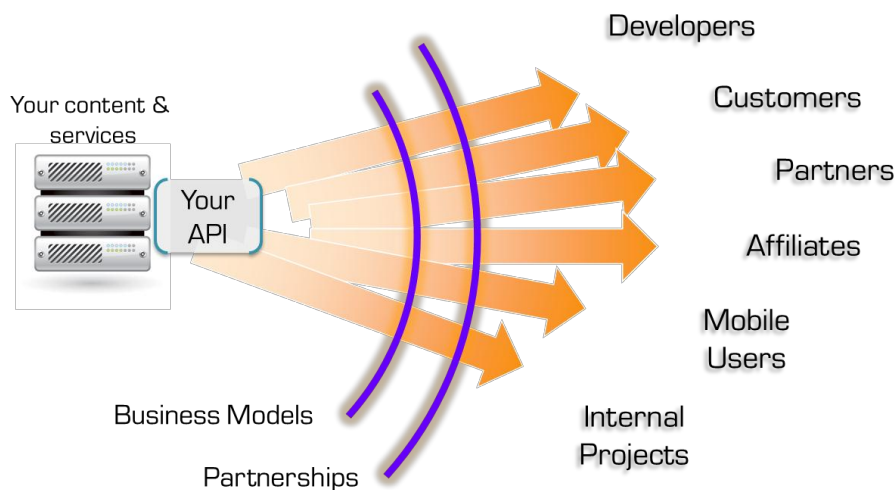
 Picture 2. Sharing API's with different stakeholders
.

2 OPEN DATA PROCESSES

Describing scenario that comes from Five Reason For scenario-Based Design John M. Carroll and has been used in a research named Innovation with open data: Essential elements of open data ecosystems. Multidimensional description will show how open data processes work and how it can be implemented in the case matkapaiva.fi.

2.1 Elements of an open government

It all starts from collecting data internally, from businesses and from individuals. Government storage data into their systems, before publicizing it they will delete or hide delicate or personal information from the files and then publish it to the web. Crucial moment before making it available is conforming and re-checking meta data is accurate, up to date and readable by the end user.

Multidimensional level in the picture 3 shows how the data is moving from government to an end user and how the data can be visualize or search with different tools. The data sets are available in XLS formats that can be machine processable or download. Machine processable means it can be read with Microsoft Excel Web App or with other applications that support XLS format. There are several different formats that publishers use.

Processing data means that the data is manipulated into a visual context, it is broadly used by researchers who are making assumptions about some situation or phenomenon. There are three different mind map areas of the processes. Tools and service area describes available APIs. Data producer and contextual level is for making sure that the data and meta data are up to date. User level is basically the final interface that the public or private sector is using and seeing.
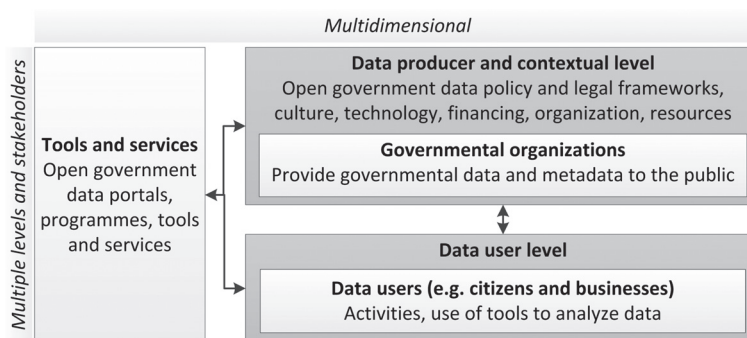
Picture 3. Elements of an open government data ecosystem derived from the literature and the scenario

2.2 Tools and services

Section about the APIs that are available in the OGD area will be mentioned in chapter 4, with practical examples. This is an area that works in two ways, the private area like, social media LinkedIn, Twitter and Facebook etc. have wide range of APIs available for the developers, that's one of the success factors of their services in online as well. Opening up your systems for outside developers can be creates a lot of value. Marketing people are using APIs to create content and update information, by their own customer relationship management systems. This is something that gives them freedom and efficient work environment to only use one system, but to create content to wider audience.

2.3 Data producer and contextual level

Making a decision about the openness of the data takes place in that scenario area. This is where data and meta data are storage, checked and analyzed. In simple words this is the information that is created about collected data. In this part different types of laws and regulations are taken into count, everything isn't public, in the public sector it can mean citizen's names or social security numbers and in the private sector it can be business secrets.

2.4 End user

When data and meta data are published to the internet, in the perfect world it should be searchable, machine readable and available without any restrictions, exceptional cases are when the data is licensed.

2 USING OPEN DATA

Internet applications are the first thing that come in to mind when thinking about possible ways to use open data (A.Poikala, P. Kola and K.A Hintikka, 2011, p 18.). Creating new business isn't the only thing you can do with open data. Open data can be used to create applications, educational purposes, research, product development, automation of processes and crowdsourcing.

Every individual or an organization can use data to it's own needs, no restrictions or limits are set. It all comes down to technical issues, evaluating the implementation, data protection, IT architecture of the government, copyrights and cooperation between different stakeholders.

3.1 Mashup

Open data applications can be called mashups, mashup is a web page or web application. Web page is a document written usually in hypertext markup language (HTML) and is accessible through browser. Web application is an application program that is stored on a remote server and delivered over the internet through a browser interface. All the data that is used in mashups are from the main data center that provides that data, so the webpage owner doesn't storage the data. In chapter 4.1 has practical examples about services that can be called mashup, they use open data to make everyday life easier.

3.2 Education research and product development

Developing new information about the information that already exists and has only a data meaning can be used in organizations and education institutions for high-quality research and to demonstrate facts. Education needs meta data, by opening up public and private data for schools makes the research more up to date and gives students a real problem to solve.

3.3 Automation

Can be used for address related information like zip code data, weather information about outdoor temperature and live updates from transportation service providers to provide better service. This is important for third parties to keep their own applications and services with current information. Automation is widely used in finance and news channels, for monitoring the surroundings.

3.4 Crowdsourcing

Crowdsourcing as a term means outsourcing information. This can mean many different things, sharing money or getting the best possible offer from the stakeholders. The best known examples of crowdsourcing are Wikipedia and Open street map (Public data an introduction to opening information resources, 2010). Wikipedia is a great example how people do the research and share it and update it. Challenge in the information sharing systems is the reliability, are the facts true. Maybe in the future when authorization will became more and more required in the internet the objectiveness in the internet will be more certain.

3.5 Commercialization

Commercialization means that an idea is modified to become a value added innovation. This has an individual, business or social wealth creation impact to an idea. When it comes to data and meta-data, it is important to take into account that value has a different meaning in the commercial context. For an individual it is important that the information that they get is relevant, by that it has to be up to date.

Businesses on the other hand try to create wealth around it, by having information that is important for stakeholders. Same aspect comes out when looking social benefit, by having an open government that gives transparent information for citizens. Chapter 4.4 will take a look into how many open data solutions are commercialized and how they have done that.

4 FINNISH GOVERNMENT OPEN DATA

This section is gathering together all Finnish government open data. It contains statistics, reports, databases and other releases that are public and available for everyone. It also includes municipality releases.

Finland has produced some high quality data resources but, in most cases, the information is published solely on PDF-format, which makes it harder to add value to data (Public data an introduction to opening information resources, 2010). Opening the data requires that laws and directives need to be changed to make that possible, this is one of the most challenging part in the processes for opening data.

4.1 Finnish transport agency (FTA)

Finnish transport agency is responsible for roads, railways, waterways and transportation systems. They also promote transportation solution development and they will open up more and more data in upcoming years.

FTA has three open data sections, digitraffic, digiroad and matka.fi. Digitraffic is for broadcasting real-time traffic and weather information. Digiroad contains Finnish national roads, street maps, bus stops, railways and crossroads. Matka.fi is the largest of them all, which stores Finnish public transportation timetables.

Digitraffic is a service, which offers real time and historical information about road weather and traffic on the Finnish main roads through Web service interfaces (Mikael Holmström, 2014, p 4). This real time means that there are weather stations that give information about the conditions on the road. Digiroad is for measuring traffic volumes and travel time. Developers have used that to show and predict how long time will they be on the road, if there is a traffic jam for example.

4.2 Finnish Meteorological institute (FMI)

The Finnish Meteorological Institute offers an online Web service interface which allows searching, browsing and downloading the weather information data sets in machine-readable format (Mikael Holmström, 2014,p 4). Data has been available and free of charge since 2013. They have 200 weather stations around Finland that produce and give live information.

FMI has a lot of data that can be downloaded for research purposes. They have real time data like, wind, temperature, humidity, atmospheric pressure, precipitin, sea level and waves data. For forecasting they have national weather surface data for 48 hours. This means that data is shared into one-hour slots for 48 hours.

4.3 Avoindata.fi

Avoindata.fi is a webpage that gather all Finnish public administration data into one place. The Ministry of Finance makes the catalogs available, this project is for transparency and the right to have an open government. This can be called as a peer-to peer platform for institutions. As of all the public administrations have their own right to release data, they can sign up and make their self-public.

When this research was made there were 1288 open data sets available. Over all content is in English, but there are also some of the data that is in Finnish only. The standard in open data is that is should be in the XLS format, not all of the data is in this form. This is for making sure that the data and meta data are machine processable. Majority of the XLS formats come from Helsinki region Infoshare, they have made the most data available.

4.4 Solutions made with open data

Data usage in web application is popular and there is competition in Finland that promotes developers to develop innovative solutions and ways to visualize OGD.

Competition name is Apps4finland, they have done that from 2009. Several great solutions and applications have been delivered for the public, from that competition. Goal is to introduce the winners and what kind of benefits they have brought with the solutions. The area of interest is weather and transportation, to be mentioned that there are other solutions from statistic to historical landmarks and maps. All solutions can be found from www.apps4finland.fi.

Commercialization that was mentioned in the previous chapter comes into play when looking into years 2010-2013. In the overall only 9% of the solutions can be called commercialized. The evaluation has been made by revenue, the organization that is behind the solution is making revenue out of data. There are three ways mentioned in that research, one is that the service is paid before usage, for example an app store, the other option is marketing revenue, having banners in the application and the third option is that data providers are paying for it. To use the service you have to pay either a single payment or a subscription from the user point of view. The third payment option, which comes from collaboration with the private sector, they have opened up an interface to be used and monitored and they pay for the analysis that they give to their customers.

Interesting part in that came out in the analysis, all commercial solutions were using location data. Location data is basically a geographical position, where you are located. This is an easy way to provide the best service for your environment, by that the solution knows where the customer is and can provide relevant and up to date information.

They all have one thing in common, they solve everyday challenges, they are not entertainment applications, and this is where social benefit is also playing a big role. Coming back to the meaning of open data it isn't just making revenue out of it, but also making public services better.

Duunitori, which is a job search engine, which combines four different interfaces where employers inform vacancies. They make revenue out of marketing, by adding banners. They also have an additional tool that helps employer to find best candidates. This service is free for the public everyone can use it to search vacant posi-

tions. Applications that are based on an earning model are available from the app stores.

Application named BlindSquare is a compass that gathers information around you. The idea behind that is that it helps blind people to get information about where they are and how far is their destination. It has unique algorithms that answer your question. It uses many different data sources to create as much information about the surroundings. It's not just in Finnish, it can also be used in 23 different languages. The only downside is that it's available only on iOs devices.

Third commercial application that is mentionable is avointuntimittaus. This is a service that uses private interfaces, but it's open for developers. The business module is that you can get real time information about your electric, water or heat consumption by hour. The developer who has made that API is making revenue from customers that use a tool to follow their consumption. This is where internet of things is used, sensors that measure consumption, more information can be found from chapter 6.3 which explains what internet of thing is.

4.4.1 Transportation solutions

Many of the solutions and apps that are made by open data are somehow related to transportation. For example Nysse, which includes Tampere, Helsinki, Oulu and Jyväskylä API's. It can be used with windows mobile phone and it's free. In the description they say it includes timetables, routes and live bus tracking. This is a great example of using different open data interfaces. Mentioned before all municipal has there own sources and API's.

Second example Ferry.fi timetable service. They have included an interactive maps solution for showing where the harbors are. Southwest Finland has a lot of islands and the demand for gathering the information about the timetables into one place was a necessity the solution provides that now. What makes the solution more complicat-

ed and valuable, they have also private sector data, not to mention Google maps for the interactive part.

4.4.2 Weather solutions

Smart solution for predicting what the weather will be like when you are on the road has been solved by using weather information. As of weather information is just a prediction, Reittisää solution shows how the weather will change between the point A and B. It shows the temperature, will it rain or is it sunny and also will the sun go down when you arrive.

As of weather information can also be used in value creating business one of the project is Sunsää. Sunsää is a platform that can change the content on a webpage, by taking into account where you are right now and what's the weather a like. For example when you are planning some activities outside the content will suggest the suitable activity that customer could do at that moment. This can be predicted for the future time with the forecast that has been made by the Finnish meteorology institution.

OGD can also be a two way street, by that I mean that information can be shared between the developers. Good example is Sujuvuusnavigaattori which is using maps and OpenTripplanner to create cycling routes. There are no open sources for cycling maps, so they have made one for sharing with the community.

5 MATKAPAIVA

This web application has been made because of the need for searching and finding the best transportation option to travel in Finland. As of Finland isn't a big country the distances are not so long, gives users option to travel with different type of transportation. Timetable search that includes all routes and stops between cities has been missing. That's why the search includes train and coach timetables.

Open data is used because of the complexity of the transportation system. The architecture and updating the data can be a resource consuming project. In the FTA interface they update it regularly and have all the information that is needed to provide the timetables. There are about 100 000 different schedules of coaches and trains that operate between cities in Finland everyday. Routes and new timetables change every two weeks automatically.

5.1 Collecting meta data

First step is to have access too the API's, this needs a license. After the license has granted, they will give a user number and a password. After the license has been approved you are ready to go. The request looks like that:

- Departure
- Destination
- Time
- Date

After the form is filled in, the engine behind sends the information to FTA interface and after few seconds the result will be available. The meta data that comes back from FTA, has more information than visualized.

The FTA interface has precise detailed information, like how much you need to walk, kilometer measurement tool and timetable related things like who is the service provider. They get their meta data from service providers, who inform FTA about the

changes and other details, they have a special interface for that as well. The transportation law regulates this.

This is a wide information packet, what Matkapäiva.fi has done it shows only the information needed, like what type of transportation should you use, and how fast you will get from point A to B.

5.2 User interface overview

Matkapaiva.fi is transportation web application that combines Finnish transport agency and Finnish meteorological institute open data. Application uses matka.fi API that gives the timetable information and FMI API for weather information. Both of the API's have a license and a special key.

Search will give the fastest option from city A to city B and the temperature in the city's A and B. Calculations between cities are made by the center of the city, in some searches it also includes traveling time inside the city of departure and arrival. This inside the city travel time is related to the system, which calculates distances from the center of the city.

Journey time inside city is included because of the change in the end result. If that would be left out, the type of transportation would be different, this is one of the challenges in the interface. They have packed everything together the train, coach, buses inside the cities and also airplane traffic information. This confuses the search the most. Picture 5 shows the weather you can see the weather information changes as well, this is possible with the weather forecast. Picture 4. Matka.fi result, from Helsinki to Tampere

1. Helsinki - Tampere				Total time: 1 h 42 min	No transfers	Walk 1.0 km
Line		Origin	Start at	Destination	Arrival at	Direction
Walk	🚶	Helsinki 🔍	15:26	Helsinki 🔍	15:30	North-east 0.3 km
Pendolino 53	🚊	Helsinki, Helsinki 🔍	**15:30**	Tampere, Tampere 🔍	17:00	North
		2 1 🍴 🚻 ♿ Service: Helsinki - Oulu				
Walk	🚶	Tampere, Tampere 🔍	17:00	Tampere 🔍	**17:10**	West 0.7 km

• Show route on the map   • Show stop listing   • Continue from destination

2. Helsinki - Tampere				Total time: 1 h 58 min	No transfers	Walk 1.0 km
Line		Origin	Start at	Destination	Arrival at	Direction
Walk	🚶	Helsinki 🔍	16:02	Helsinki 🔍	16:06	North-east 0.3 km
InterCity 55	🚊	Helsinki, Helsinki 🔍	**16:06**	Tampere, Tampere 🔍	17:52	North
		2 1 🚻 A ♿ 🚶 🚲 Service: Helsinki - Oulu				
Walk	🚶	Tampere, Tampere 🔍	17:52	Tampere 🔍	**18:02**	West 0.7 km

• Show route on the map   • Show stop listing   • Continue from destination

3. Helsinki - Tampere				Total time: 1 h 42 min	No transfers	Walk 1.0 km
Line		Origin	Start at	Destination	Arrival at	Direction
Walk	🚶	Helsinki 🔍	16:26	Helsinki 🔍	16:30	North-east 0.3 km
Pendolino 91	🚊	Helsinki, Helsinki 🔍	**16:30**	Tampere, Tampere 🔍	18:00	North
		2 1 🍴 🚻 ♿ Service: Helsinki - Pieksämäki				
Walk	🚶	Tampere, Tampere 🔍	18:00	Tampere 🔍	**18:10**	West 0.7 km

• Show route on the map   • Show stop listing   • Continue from destination

Picture 4. Matka.fi result, from Helsinki to Tampere

Comparing picture 4 and picture 5, which are from matka.fi and matkapäiva.fi, the biggest difference is that matka.fi provides a lot more information than matkapäiva.fi. The reason for that is manipulation and deleting information that is in the development plan that matkapäiva.fi has. As mentioned before the idea is to get from point A to B as fast as possible and with no walking or map information. They both use the same square method to visualize the data in order. Picture visualization is better in the matka.fi case, which has icons for walking and transportation type.

Picture 5. Matkapaiva.fi/search/result, between cities Helsinki and Tampere

5.3 How to make a search

Search engine that can be found from the address www.matkapaiva.fi, has 4 slots that can see from picture 6.Things that you have to type in are, place from where the departure happens and where are you going, date when the departure is happening and the time with options. List of cities that are available is limited. List of cities that are can be searched are in the reference part and are referenced by available tags.


Picture 6. www.matkapaiva.fi main page, 4 slots and one activation button for search

Date slot is visualized by calendar view, which allows changing the date by one month ahead. Date is limited because of the available meta data. Calendar view won't allow to choose dates that aren't available. Dates that aren't available are coloured light grey. Picture 7 shows the grey are which dates aren't available. Calendar option is meant for planning journey, otherwise the date is fixed for the date that the website is used. Typing the date to the slot is not made possible to make the search faster.
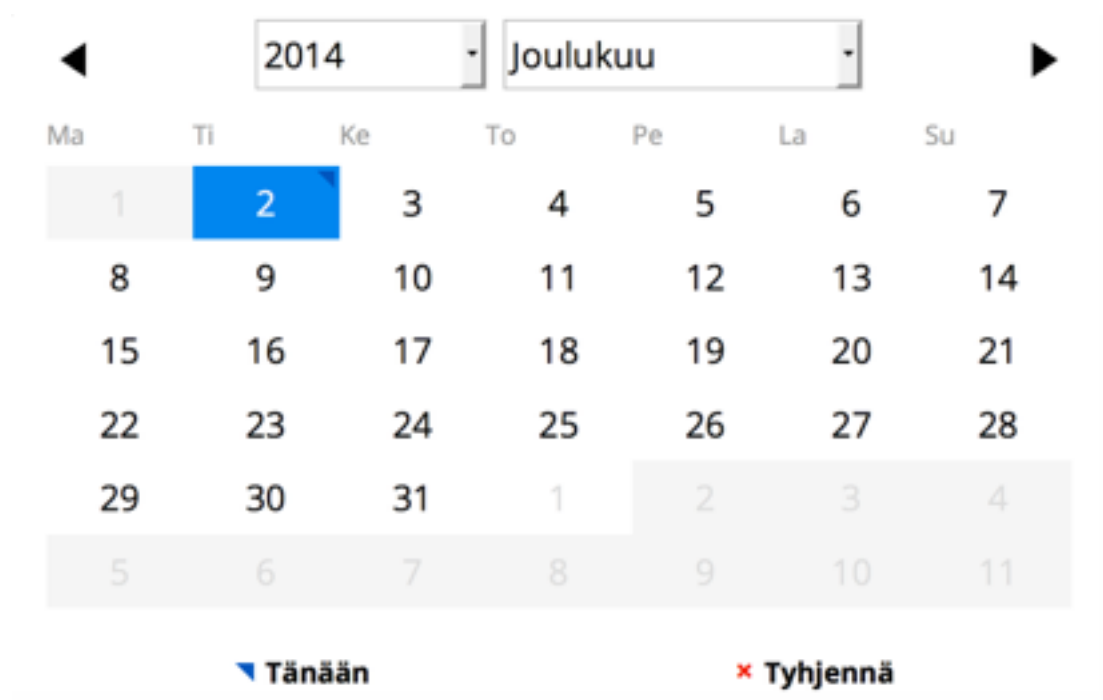
Picture 7. Calendar view from matkapäiva.fi, for choosing the date of departure

Time slot shows the time when the departure will happen, it is visualized in quarters. This fan content view, that gives a clear time frame when the departure happens. By selecting the right time, selected time will be became blue, example can be found from picture 8. After slots have been filled and search button has been clicked search for the timetables will start. The actions behind were explained before.

09:00
09:15
09:30
09:45
10:00
10:15
10:30
10:45
11:00
11:15
11:30
11:45
12:00
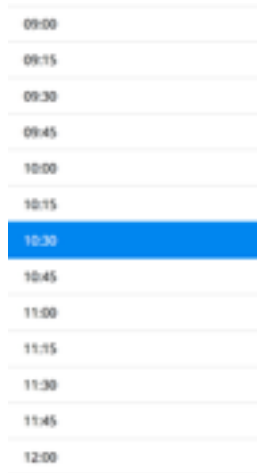
Picture 8. Time fan for choosing the departure time

5.4 Search result

Search result will show three options that are fastest. It also shows the date, time, transportation type, journey duration and outside temperature. All timetables are estimations and as of the data is from third party, it is impossible to promise that the transportation will take you from point A to B. Temperature is also an estimation if the search is made for future and it usually is, for example tomorrow at 2 pm. Picture 9 shows where are you going from and to, date, with time of departure and arrival, temperature in both city's, summery and transportation type. In some cases when exchange is needed, it also shows where the exchange happens.

Picture 9. Search result form matkapaiva.fi, showing the result between Helsinki and Turku

5.5 Value creation

The uniqueness of the service and search is that not all the APIs content is used. By cleaning up the meta data, that is available, creates more value to the end user. Value that is created is not wealth value, but social value. Having the possibility to get the fastest transportation option can be a crucial thing when planning a journey. Mentioned in the previous chapters open data usage doesn't have to be commercialized, it can also mean providing better solutions for the public.

5.6 Audience overview

The website has been public since June 2014. For testing and developing purposes, no marketing has been done for driving traffic. The tool that is used in the case is Google analytics. Google analytics is a free statistics tool that keeps track on incoming traffic. It only shows visitors that have allowed JavaScript. This is a great tool that allows website owner to know where the customers are coming from, what pages they are using, how long they will stay on the page and what device they are using. These are crucial things when developing and provide a service to you customers.

5.6.1 Overlook of the data

Data range for this analysis is between June 2014 and January 2015. The overall numbers are quit low. Unique visits together 834 and page over all sessions 1146. These two are the most valuable measurement numbers. This shows that there are 312 returning users, that is a bit lower than half of the users.

This can be explained by saying that users didn't found the service useful enough to return. The most interesting data, that analysis shows for the developer is how many searches were made? This number can be found under pages views that were 301. From the over all page views it's 26,27%, one out of every 4 visitors made the search and got the result. As the picture 10 shows most of the visitors just landed the page, with out making the search, 520 of them.

Page		Pageviews	% Pageviews
1. /		520	45.38%
2. /search/results/		301	26.27%
3. /info/about/		137	11.95%
4. /info/contact/		83	7.24%
5. /frontpage/view		54	4.71%

Picture 10. Google analytics behavior, which sites have been used

The average time on page is 1 minute and 7 seconds. This can be explained by the feedback that was coming from the test group. The first version was made as a flow chart and it didn't explain the route clearly and the transportation type wasn't available.

5.6.2 Viewers segment

Looking into segments where the customers are coming from, by location, behavior, how long time they were on the site. Other things that analytics show are engagement and new vs. returning and also from the technology angle like what browser was used and was it a mobile or a desktop user. Some of the results were explained before and now let's combine them.

Majority visits were made from Finland, top three cities are Helsinki, Pori and Tampere, picture 11 has a list of top visited cities. Approximately 90 % of the visits were made from these three cities. Engagement level was low from 0-10 seconds, over all it was clicking or occasion that they were on the site. Practically all of the users have landed the main page and directly, by knowing the web page name.

1.	Helsinki	104 (29.21%)
2.	Pori	101 (28.37%)
3.	Tampere	93 (26.12%)
4.	Turku	14 (3.93%)
5.	Espoo	11 (3.09%)
6.	Oulu	8 (2.25%)
7.	Vantaa	4 (1.12%)

Picture 11. Google analytics location, from where viewers have visited matkapaiva.fi

Taking into account that there are many different browser options, the developers have to take that into account, when developing a web page. Top three browsers that visit are Chrome, Firefox and Safari. Great user experience that has been built for the site is that when the browser's size is reduced the site will change its form. This enables mobile and tablet users to make the search and reading better. From the statistics there were 338 desktop-, 73 mobile-and 48 tablet users.

These statistics can be used for making a customer analysis. From the statistics, we know where they are, what is the device that they are using and how they got to the site. All of the statistics are available as appendices. This statistics isn't reliable because the period in hand has been a testing period and most of the visits are just feedback. For making a better conclusion about the costumers it is needed to drive more traffic on the site and have a longer period of time for the users to get to know the service.

5.7 Improvements

This period of time when the application has been public for the test group, it would be wise to mention that some parts of the meta data should be changed. The challenge is that there are too many different coach companies that provide service around Finland. The meta data should be tagged better, by that I mean all the service providers should be tagged. Right now the meta data has a lot of information that isn't relevant or important for the user. Example when the search is made from Helsinki to Tampere, it offers that about the service provider:

" *pika, Liikennöitsijä VÄINÖ PAUNU OY REITTI-555541 937644/154584/793331, Service: Kamppi, kaukoliik. – Virrat"*

 This information is not useful, because it's a full tag. In the case where FTA would change the tag into *VÄINÖ PAUNU OY* it would be possible to convert the tag into Matkahuolto. The same thing should be done with all of the transportation providers and manually. This is just an example of the tag that is right now available from the

FTA API, some of the tags are more complicated. One reason for that is that they only have the obligation to add everything, which is related to any specific route between cities.

With train providers it's easy because of the monopoly, only one service provider. The tag below is from the train timetable point of view.

"2 1 🐎 A ♿ 👫 🚲 🚂 Service: Helsinki – Tampere"

Little pictures show additional services that are provided on the train. No converting is needed in that case, because the only service provider is VR.

Open data developer's point of view I would say that the change in the API should be made. By tagging all the service providers in to one data set, would also make the linking to the web sites better.

6 SIMILAR SERVICES

Long distance timetables are something that is not available in many websites. They can be found from the service providers web pages, which are the most reliable information sources and also from the Finnish Transport Agency search engine, called matka.fi. New website that has also the timetables and prices is called pikavuorot.fi.

6.1 Service providers

Finnish long distance transport system has been closed for years. After 2011 some of the routes have been open for private companies that can provide transportation services between cities. Nationwide transportation will be open in 2019. Now there are 4 service providers, Matkahuolto, Onnibus,Onniexpress and Vr. First three are coach providers and Vr is providing train transportation. There are also smaller coach providers, the ticket sales are run by Matkahuolto. Smaller ones don't sell tickets but they have websites about the routes and services. I have combined them under Matkahuolto. Timetables can be search form companies home webpages. This is time consuming, but I will give the most reliable information and also a possibility to purchases tickets.

July 2014 a new web service was launched that has coach and train ticket prices. Pikavuorot.fi has made a promising solution in that market area. They provide the most information about the timetables and also the price, that is important for making a decision. They are not using FTA API, they use information straight from the providers websites. This is something that is mentioned previously as well, were private companies have opened up their own APIs.

6.2 Analysis

As of the case in hand uses open government open data, this analysis will not include, services providers applications or Pikavuorot.fi. Leaving Pikavuoro.fi out because of the meta data that they use, as mentioned before they use data from the service providers webpage.

This analysis will take a look at the Matka.fi application and Matkapaiva.fi. Three practical elements that are relevant in web design: Design, how the data is visualized, content like text and the forms that it is displayed and actions how easy it is to use the application.

6.2.1 Design

Data and meta data are pieces of information, the value creation comes when it's visualized in a smart way. Form meta data point of view both applications use same data sources. Difference comes from visualizing, matka.fi gives out more details, like map view, specific address option, walking distance calculation and specific explanation about the service provider. Matkapaiva.fi only gives information about the total time of the travel and type of transportation.

6.2.2 Content

In content view it is important to mention that the more information is visualized the better. So Matka.fi has a lot of specific information available, that makes the data usage in the searcher prospective most valuable. On the other hand too much information can complex the decision making. Illustrational pint of view Matkapaiva has a clean content that doesn't complicate the end result. Usage of content isn't used perfectly in Matkapaiva case. Improvement section will take a look into what can be made in that area.

6.2.3 Action

Search to work, it is required that the users types in three key slots:

- From
- To
- Time
- Date

Both of the services have that, this is more of a minimum, which is needed. Additionally, matka.fi has more to make the search, from the map and from the index list. It's not enough if you just type in the city name, either way it is needed to choose from the list that they give. Matkapaiva.fi on the other hand only suggests the city.

Too many different options are not convenient when it comes to finding transportation between two cities. From the action point of view matkapaiva.fi has smarter options in the form filling. In case where you search for Tampere only one option, but making a search with matka.fi picture 13, you will get 18 different options that are listed in picture 12. This can be frustrating, comparing this to matkapaiva.fi, we can see that there are just one option. This makes the search easier from the searchers prospective, but if the searcher is looking for more specific destination, 18 options can be more useful.
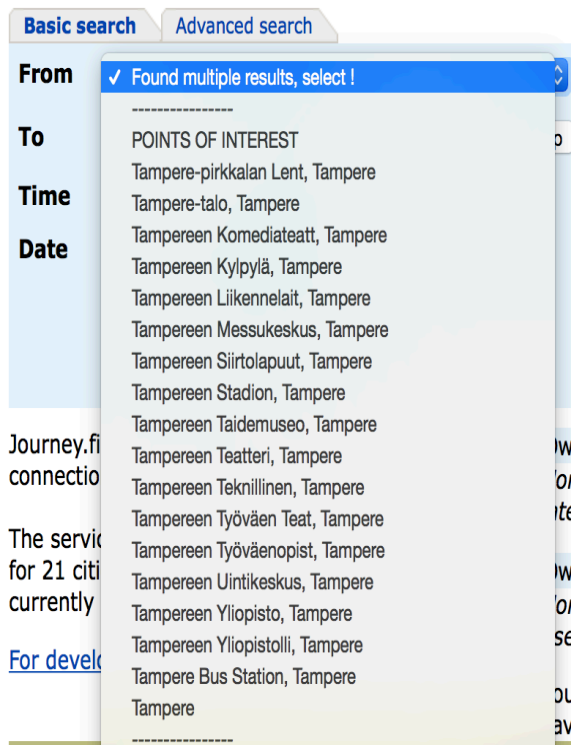
Picture 12. Matka.fi, selecting a specific location, suggestion fan
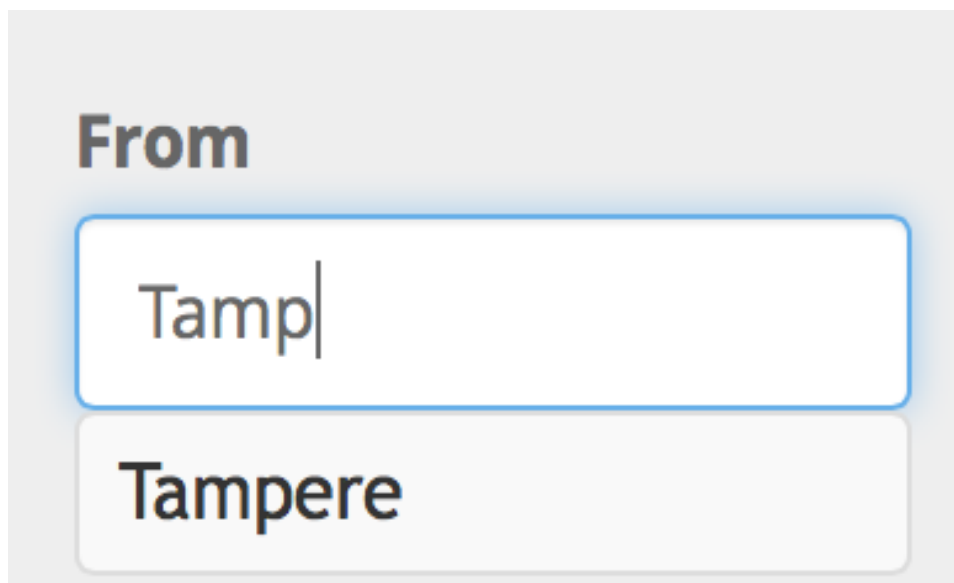
Picture 13 Searching for Tampere, suggests the best possible option, matkapaiva.fi

6.2 Internet of things

This chapter is added as additional information what will future bring in the area of connecting things and open data. To provide real time information, devices and data has to be connected with each other. There are cities in Finland, like Helsinki, Tampere and Oulu, which have tracking devices inside the busses that operate in the city. VR has also that kind of service. These kinds of real time monitoring systems are more and more popular around the world as well.

These sensors that are connected with internet are called Internet of Things. They have opened up the interfaces for developers as well. As of more and more devices will be connected with each other, the possibility to create better services is also rising. In fact only in Tampere there are over 30 different applications created from bus timetables inside the city. No all of them are free, but it shows that business can created solutions from tracking data as well.

This new way to sensor and monitor our surrounding is been funded by the European Commission who has give 8 million euros for European universities and telecom companies to install some 18 000 sensors, which keep track on air pollution, noise and other pollution factors that affect peoples lives in big cities.

There is lot to learn when it comes to our well being in the environment that we are living in and the public services that we are using.

Internet of things will be a wider topic in up coming years. This is not just a topic of open data it also has an effect and impact to several different industries. It is widely used in the financial and media industry to monitor information flow from different sources. The criticism that they have gotten with sensors and monitoring is privacy concerns. Can citizens count on the private and public sector that their just monitoring the over all, not the individuals. This is something that needs to be researched. Coming back to the chapter about elements of open data, it is highly important to delete or hide delicate or personal information from the flies that are published in the web.

CONCLUSION

The vision of this thesis was to introduce open data and how developers can innovate and provide commercial and public services. This thesis can be used, as an introduction to open data, there is so much to learn and develop in that area. Developing a timetable comparison application with open data was also done and explained. The hypothesis is, it possible to create applications with open data and what should be taken into a count before using open APIs. The theory about open government data which says, that licensing and open distribution of data and meta data, will create more services has been demonstrated with a practical application.

Open data is a new thing in the public government sector, literature and research papers that are been used in this thesis are up to date and not older than 5 years. This means that the research that was made is needed to show practical examples and explain what can be done in the area. Providing services that include several sources, like Matkapäiva.fi are needed and there is a market gap to be filled. Not only is open data a new thing, there are more than 700 different solutions in Finland, interactive maps and everyday applications that are using different types of sources. Finland can be called as a pioneer in the developing new solutions out of open data. It is mentionable that the trend is growing around the world.

Beneficially talking open data benefits are: transparency, applications for the citizens and optimization of administrative processes. Having the possibility to see how the decisions have been made and what is behind the solution is a powerful tool for the public to see how tax money is used. The solutions side, which this thesis analysis, clears out that the more services are developed out of open data the more, we have services and social value created. It is mentionable that only tiny part of all the appli-

cations and visualizations can create revenue. This shouldn't be a boundary for developers who use open data. Finding creative ways to take advantage out of open data is promoted by the European Union.

The European Union announced at the end of 2014, that they support open data across Europe with 14.4 million euros. This money will be shared with research and innovation. This means that opportunity to develop and find funding for open data solutions is easier. 7.8 million euros will be given to Open Data Incubators for Europe, to create economic and social value. Researcher will get 3.7 million euros to find answers like " how can web data be used".

Further research could be about how private companies could benefit from opening up their data. One of the examples that I mentioned in the paragraph 4.4 which is using private data and making revenue out of it. This is an interesting and research needed area. Private companies are the ones that have sensors for collecting information about our surroundings. Most of the information that is available right now is related to media and statistics, which can only be benefited for a short time. I would be great if telecommunication companies would open up global positioning system data. For examples where have I been and what data has been collected about me. This information can be asked form the telecommunication companies, but they don't have any API's for that so they will give that information on paper. This is the part where machine-processable comes into play. Another area of business, which is retail, they have a lot of data. Different customer promotion cards, which collect data about our shopping behavior and can go back many years, could be used better. There is a small amount that they really use themselves.

REFERENCES

A. Poikola, P. Kola and K.A. Hintikka, Public data. An introduction to opening in-formation sources, Ministry of Trans- port and Communications, Helsinki, Finland 2011

Mikael Holmström, Open data basedRWIS for mobile devices, Helsinki, Finland Jun 2014

Maija-Liisa honkola, Nina Kukkurainen, Lea Saukkonen, Anu Petäjä, Janna Ka-rasjärvi, Tarja Riihisaari, Roope Tervo, Mikko Visa, Juhana Hyrkkänen, Reija Ruuhela, The Finnish Meteorological Institute - Final Report For the Open Data Pro-ject

House of Representatives Standing Committee on Science and Innovation, Chapter 2: Innovation and Commercialisation - Concepts, Definitions and Metrics, 2006

What is an API, 3scale USA, 6 June 2012
http://www.3scale.net/wp-content/uploads/2012/06/What-is-an-API-1.0.pdf

Anneke Zuiderwijk, Marijn Janssen and Chris Davis, Innovation with open data: Es-sential elements of open data ecosystems, Faculty of Technology, Policy and Man-agement, Delft University of technology, Delft, The Netherlands, 2014

Samuel Gibbs, 2014. EU commits €14.4m to support open data across Europe. The guardian 04.11.2014. http://www.theguardian.com

FIGURES

Picture 1. Acclar blog, Is open public data a service or a right? Referred 20.02.2015.
http://acclar.org

Picture 2. Website of 3scale Inc. What is an API? Referred 06.06.2012.
http://www.3scale.net

Picture 3. Anneke Zuiderwijk, Marijn Janssen and Chris Davis, Innovation with open
data: Essential elements of open data ecosystems, Faculty of Technology, Policy and
Management, Delft University of technology, Delft, The Netherlands, 2014

Picture 4. Website of Finnish Transport Agency. Referred 12.02.2015.
http://www.matka.fi.

Picture 5. Website of Kaarel Kõrvemaa. Referred 20.02.2015.
http://www.matkapaiva.fi

Picture 6. Website of Kaarel Kõrvemaa. Referred 20.02.2015.
http://www.matkapaiva.fi

Picture 7. Website of Kaarel Kõrvemaa. Referred 20.02.2015.
http://www.matkapaiva.fi

Picture 8. Website of Kaarel Kõrvemaa. Referred 20.02.2015.
http://www.matkapaiva.fi

Picture 9. Website of Kaarel Kõrvemaa. Referred 20.02.2015.
http://www.matkapaiva.fi

Picture 10. Website of Google Inc./ analytics. Referred 28.01.2015.
www.google.com

Picture 11. Website of Google Inc./ analytics. Referred 28.01.2015.
www.google.com

Picture 12. Website of Finnish Transport Agency. Referred 12.02.2015.
http://www.matka.fi.

Picture 13. Picture 9. Website of Kaarel Kõrvemaa. Referred 20.02.2015.
http://www.matkapaiva.fi

Available Tags

Akaa','Alajärvi','Alavieska','Alavus','Asikkala','Askola','Aura','Brändö','Eckerö','Enon
koski','Enontekiö','Espoo','Eura','Eurajoki','Evijärvi','Finström','Forssa','Föglö','Geta','
Haapajä-

jä-

rvi','Haapavesi','Hailuoto','Halsua','Hamina','Hammarland','Hankasalmi','Hanko','Harj
av-

alta','Hartola','Hattula','Hausjärvi','Heinola','Heinävesi','Helsinki','Hirvensalmi','Hollol
a','Honkajoki','Huittinen','Humppila','Hyrynsalmi','Hyvinkää','Hämeenkoski','Hämeen
kyrö','Hämeenlinna','Ii','Iisalmi','Iitti','Ikaalinen','Ilmajoki','Ilomantsi','Imatra','Inari','I
nkoo','Isojoki','Isokyrö','Jalasjärvi','Janakkala','Joensuu','Jokioinen','Jomala','Joroinen'
,'Joutsa','Juankoski','Juuka','Juupajoki','Juva','Jyväskylä','Jämijärvi','Jämsä','Järvenpää'
,'Kaarina','Kaavi','Kajaani','Kalajoki','Kangasala','Kangasniemi','Kankaanpää','Kanno
nkoski','Kannus','Karijoki','Karkkila','Karstula','Karvia','Kaskinen','Kauhajoki','Kauha
va','Kauniainen','Kaustinen','Keitele','Kemi','Kemijärvi','Keminmaa','Kemiönsaari','K
em-

pele','Kerava','Keuruu','Kihniö','Kinnula','Kirkkonummi','Kitee','Kittilä','Kiuruvesi','K
ivijärvi','Kokemäki','Kokkola','Kolari','Konnevesi','Kontiolahti','Korsnäs','Koski','Tl','
Kotka','Kouvola','Kristiinankaupunki','Kruunupyy','Kuhmo','Kuhmoinen','Kumlinge','
Kuo-

pio','Kuortane','Kurikka','Kustavi','Kuusamo','Kyyjärvi','Kärkölä','Kärsämäki','Kökar',
'Köyliö','Lahti','Laihia','Laitila','Lapinjärvi','Lapinlahti','Lappajärvi','Lappeenranta','La
pua','Laukaa','Lavia','Lemi','Lemland','Lempäälä','Leppävirta','Lestijärvi','Lieksa','Liet
o','Liminka','Liperi','Lohja','Loimaa','Loppi','Loviisa','Luhanka','Lumijoki','Lumparlan
d','Luoto','Luumäki','Luvia','Maalahti','Maaninka','Maarianhamina','Marttila','Masku','
Merijärvi','Merikarvia','Miehikkälä','Mikkeli','Muhos','Multia','Muonio','Mustasaari','
Muurame','Mynämäki','Myrskylä','Mäntsälä','Mänttä-

Vilppu-

la','Mäntyharju','Naantali','Nakkila','Nastola','Nivala','Nokia','Nousiainen','Nurmes','N
ur-

mijärvi','Närpiö','Orimattila','Oripää','Orivesi','Oulainen','Oulu','Outokumpu','Padasjo
ki','Paimio','Paltamo','Parainen','Parikkala','Parkano','Pedersören
kunta','Pelkosenniemi','Pello','Perho','Pertunmaa','Petäjävesi','Pieksämäki','Pielavesi','
Pietar-
saari','Pihtipudas','Pirkkala','Polvijärvi','Pomarkku','Pori','Pornainen','Porvoo','Posio','
Pudasjä-
jä-
rvi','Pukkila','Punkalaidun','Puolanka','Puumala','Pyhtää','Pyhäjoki','Pyhäjärvi','Pyhänt
ä','Pyhäranta','Pälkäne','Pöytyä','Raahe','Raasepori','Raisio','Rantasalmi','Ranua','Rau
ma','Rautalampi','Rautavaara','Rautjärvi','Reisjärvi','Riihimäki','Ristijärvi','Rovaniemi'
,'Ruokolahti','Ruovesi','Rusko','Rääkkylä','Saarijärvi','Salla','Salo','Saltvik','Sastamala'
,'Sauvo','Savitaipale','Savonlinna','Savukoski','Seinäjoki','Sievi','Siikainen','Siikajoki','
Siikalat-
va','Siilinjärvi','Simo','Sipoo','Siuntio','Sodankylä','Soini','Somero','Sonkajärvi','Sotka
mo','Sottunga','Sulkava','Sund','Suomussalmi','Suonenjoki','Sysmä','Säkylä','Taipalsaa
ri','Taivalkoski','Taivassalo','Tammela','Tampere','Tarvasjoki','Tervo','Tervola','Teuva
','Tohmajärvi','Toholampi','Toivakka','Tornio','Turku','Tuusniemi','Tuusula','Tyrnävä',
'Ulvi-
vi-
la','Urjala','Utajärvi','Utsjoki','Uurainen','Uusikaarlepyy','Uusikaupunki','Vaala','Vaas
a','Valkeakoski','Valtimo','Vantaa','Varkaus','Vehmaa','Vesanto','Vesilahti','Veteli','Vi
ere-
mä','Vihti','Viitasaari','Vimpeli','Virolahti','Virrat','Vårdö','Vöyri','Ylitornio','Yliviesk
a','Ylöjärvi','Ypäjä','Ähtäri','Äänekoski'