



INNOVATION AND ENTREPRENEURSHIP THROUGH OPEN DATA

INTRODUCTION

Publishing data from local government is useful by itself. However, the real power of the data and what it can tell you about a local area is most likely to be seen when other organisations develop tools and applications which present the data in a way which can be easily understood.

Once data is made available, it is open to businesses (often small and medium enterprises or SMEs) to analyse and create business from selling that analysis, or to develop and sell tools making use of the data. Such innovation and entrepreneurship can act as a driver in the improvement of local government services, with little or no cost to an administration other than that of releasing the data for public consumption.

Government in many partner countries is seeking to promote innovation by organising activities such as hackathons, developing case studies and by encouraging local government to increase the amount of data made available.

Some examples of how open data has been or could be used in innovative ways through entrepreneurship in partner countries and municipalities are described in the rest of this document.

HOW TO RUN A SEMINAR TO IDENTIFY OPEN DATA THAT INTERESTS BUSINESSES AND POLITICIANS

Karlstad municipality in Sweden has found the main problem is that people don't know what open data is. They don't know why data should be published as open data and find the whole topic too overwhelming to even talk about.

A further problem arises from the fact that data is owned by more than one administration, so the decision to publish may have to be taken in several steps. Both politicians and civil servants need to be on board if the data is to be released.

In a municipality, the main reason to open up data is not to make money, it is to create a transparent and more open local government. Of course, businesses will use the data and help a municipality to both create better services and make money, but that is a secondary benefit, albeit one which may help convince people to put money into open data rather than directly into improved health care or better schools. The open data can help to achieve both of these objectives without the direct spend.

A representative from Karlstad says: "During the year I had meetings with the Värmland county administrative board who were experiencing the same problems in their project, Digital Baltic (Interreg Baltic Sea), so we decided to work together. I spoke at their first meeting in Stockholm and the main idea was to have workshops with local and national governmental organisations in our regions. "We created a day that started out with talking about what open data is, what is open and what has already been done in Sweden.

"In the afternoon we had an 'open space' and the main question was: what is open data and how can



we use open data in our organization? Everyone could ask their own question and discuss it with the others. There were some key words that followed us through most answers: working together, transparency and using data within our organisation.”

Digital Baltic is a project co-funded by the Swedish Institute. It consists of partners from Sweden, Estonia, Lithuania, Norway, Poland and Denmark, and aims at promoting digital innovations in Public Sector Information (PSI)/Open Data from a citizen perspective in the Baltic Sea Region.

Today, PSI re-use and Open Data in Europe is strongly underutilised and the economic value is high. Hence, there is huge potential for the Baltic Sea Region to spur innovation within this field, especially for SMEs. In order to overcome the challenge of driving towards a digital single market, there is a need to stimulate digital innovations within the re-use of PSI and Open Data.

The Digital Baltic project aims to meet this need by involving citizens and businesses through the working methods of Open Space in each partner region. The results from the regional Open Space meetings, 20 digital innovations in PSI/Open Data, will result in a preliminary study. This study will provide solid ground for a project application in the Baltic Sea Region in the near future. The results of the project will also be one step in the work towards a digital single market in the Baltic Sea Region.

The results from the Open Space workshop were not direct new innovations – instead, the focus was on regional co-operation to open up data to the region’s citizens. The results from the workshop were presented to the directors of public sector

organisations in Värmland on 27 August. Hopefully this workshop was a start to making Värmland more open towards publishing open data and PSI – which will lead to more innovations in the field.

Open Space Technology is one way to enable all kinds of people, in any kind of organisation, to create inspired meetings and events. Over the last 20+ years, it has also become clear that opening space, as an intentional leadership practice, can create inspired organisations, where ordinary people work together to create extraordinary results on a regular basis.

In Open Space meetings, events and organisations, participants create and manage their own agenda of parallel working sessions around a central theme of strategic importance, such as: What is the strategy, group, organisation or community that all stakeholders can support and work together to create?

With groups of 5 to 2000+ people — working in one-day workshops, three-day conferences, or the regular weekly staff meeting — the common result is a powerful, effective connecting and strengthening of what’s already happening in the organization: planning and action, learning and doing, passion and responsibility, participation and performance. Examples from the Digital Baltic project of good practice in the use of open data can be found at: <http://digitalbaltic.eu/good-practices-on-psi-open-data/>

Digital Baltic is also running a seminar with local businesses (SMEs) in November 2014 to promote and stimulate ideas around the use of PSI and open data and gain an understanding of what may interest SMEs in the area of open data.



WHAT DOES OPENING PUBLIC DATA MEAN FOR CITIZENS

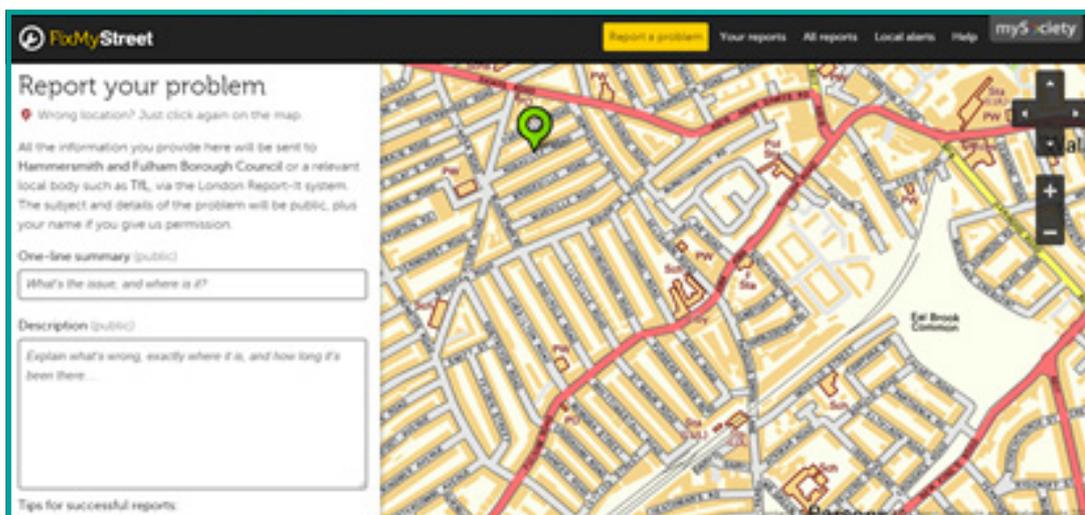
In a municipality, the main reason to open up data is not to make money, but to create a transparent and more open local government. Of course businesses will use the data and help us create better services and also make money for the municipality (by selling the data), but that is a secondary benefit. One of the biggest problems is convincing decision makers to put money into open data rather than improving health care or better schools.

Historically, the local citizen had to request data from their local administration. This led to a culture of believing that the public sector was secretive and

data in a way that can be accessed and analyzed can act as an incentive for public organisations to improve their performance.

However, the benefits of open data go far beyond this. The public sector generates data about many topics of interest to the citizen in their daily lives – traffic, flood risk, crime and health are just a few examples.

Whilst the average citizen is not going to want to look at this as a raw dataset, once the data is used by entrepreneurs, technology companies and individuals to design and develop mobile apps or tools then the data becomes of great interest and potential benefit to the public.



not fully accountable to the public. Publishing open data has, at least in theory, made it easier for the citizen to find out how their council is performing and how they are spending their money – so increasing transparency and accountability. Publication of

For example, in the UK the 'FixMyStreet' web site or mobile application allows a member of the public to report a problem at a location identified by postcode, street name or area or by geolocation and then keying in details of the problem. The application also shows existing reports on the map so that



HOW STANDARDS HELP

Local open data is most useful when like datasets from different localities can be identified, compared and combined. This is complicated where power is devolved down to regional and local government, as consistency of open datasets cannot be mandated across a country or the EU as a whole. However mechanisms can be implemented to aid local governments in aligning where they see public and mutual advantage in doing so.

For some time now, central government in many countries has made data more available for consumption by other organisations and individuals, in both public and private sectors. As technology and understanding of data publishing become better understood, a demand has emerged for the same approach by local government.

For a business, there is not necessarily a creditable case for spending time and resources developing a tool or application which will only work with published open data of one municipality. Where the data is localised to that extent, the audience is also likely to be only local, thus severely limiting the potential for a business to make money from the tool. However, if the data is published to standards that have been agreed nationally – or even better internationally – then the data has more potential. Even a simple application which uses a single dataset – for example, the location of public toilets – has a much bigger potential customer base if it works anywhere in a country rather than just a small municipality.

Once an application becomes more sophisticated, even allowing for comparison between municipalities, or shows data across multiple municipalities, then the potential business that can be built around that application is much more appealing.

However, in order to achieve this there are standards which need to be defined which both:

- Define the semantics for each dataset – what everything referenced in the dataset is called. For example, a standard name for a municipality, a status type for a licence, an identifier for a service and so on.
- Define the format of the data – what fields a dataset contains and what is an acceptable entry for each field (from a defined list, can it be blank etc). The data format is defined by a schema. Datasets from different sources which match the schema definition can typically be aggregated and compared.

If open data is published according to a schema and this is referenced for anyone looking at the data, then anyone using the data will be able to link that dataset to others published using the same schema and build an application or tool which makes use of the linked data from various sources.



SEARCHING WITHIN IMAGES

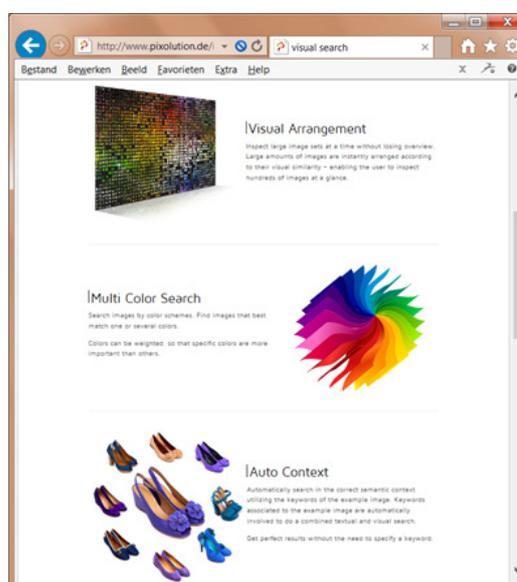
Pixolution is a German search engine opening up images on the worldwide web in a different way from other search engines. It was developed as a startup in 2009 by Professor Dr. Kai Barthel and his students at the THW Hochschule für Technik und Wirtschaft in Berlin. Professor Barthel's main areas of interest are image processing, visual and semantic image searching.

THW is Berlin's largest University of Applied Sciences offering around 70 bachelor and master courses in engineering, economics, information technology, culture and design. At this moment Pixolution is an independent and privately held company.

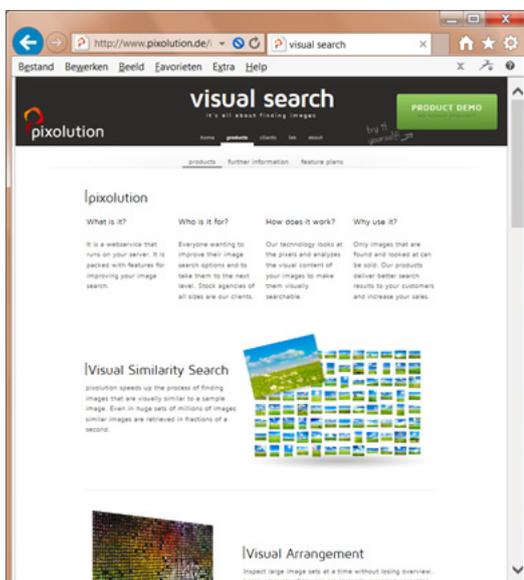
Pixolution uses different techniques to search and retrieve images in comparison to standard search engines. It looks at the image pixels and analyses the content of a picture in order to make it searchable. It is also available as a plugin for the enterprise search server Apache Solr (open source) whereby the text search capability of Apache Solr is enhanced by integrating visual image search services.

The result of this combined approach is that the search engine finds images that match both the way the images look and the content and context described by the keywords linked to the image. The advantages are that it finds images similar to a sample image (even in a very large collection of images) and that pictures with a specific colour combination can be looked up.

Although Pixolution, at this moment, mainly focuses on finding clients in the e-commerce market and photo stock field, it is a very interesting tool for museums, archives and other cultural heritage organisations for opening up historical images. Most search engines and websites present their search results with images in an unsorted way and web users are not able to view more than 20 to 50 pictures at one moment. Research also suggests that consumers only look at the first two or three results pages, so relevant pictures are not always the ones shown.



Based on their main technique, Pixolution developed several products such as Visual Similarity Search, Visual Arrangement, Multi Color Search, Auto Context Search and Black and White Filter.



For further details see Pixolution (<http://www.pixolution.de>).

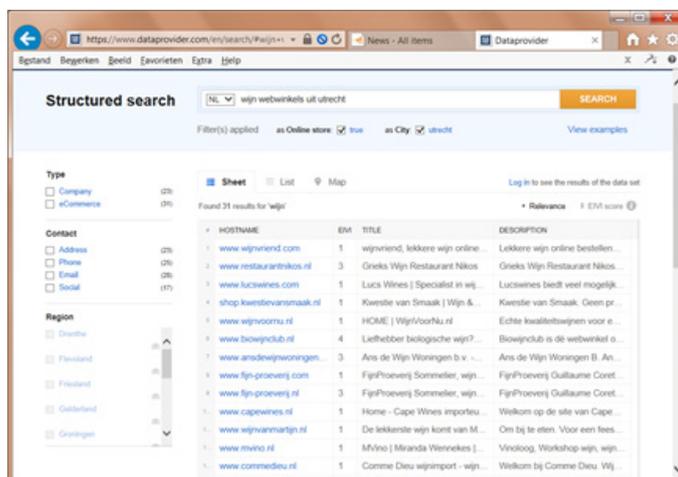
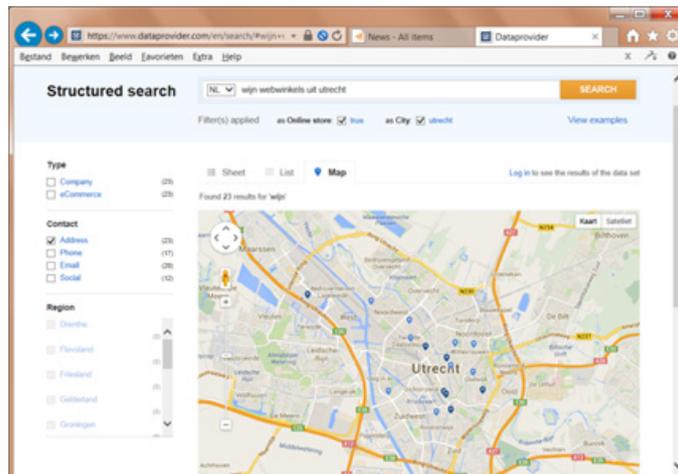


CRAWLING THE WEB AND CREATING INFORMATION FROM OPEN DATA

The internet is often regarded as a chaotic collection of data where everybody can contribute information, search engines offer bad search options and endless pages of results where information of good quality is hard to find. Several research projects have already showed that web users have difficulty in defining search strings and only look at the first two to three results pages. Designers of web pages are using several techniques in order to get a high ranking in the result scores of search engines and consumers often find commercial web pages and data/information that is not very useful.

Dataprovider is a Dutch based company founded in 2008 that crawls the web and websites for data and presents this in a structured way, different from normal search engines. It simply provides business consumers an overview of web data that traditional search engines don't offer.

Dataprovider aims to structure the crawled web data and provide business clients with the information they need for daily business, staying competitive, research and finding new business opportunities. The data is structured in such a way that customers of Dataprovider can make their own selection of data. For instance, you could search for websites on iPhones or for all Chinese restaurants in England, then start to filter these to find more specific information, as can be seen in the company promotion video: <http://vimeo.com/42135451>. The example below is created with the search for all online shops in Utrecht, Netherlands selling wine. The data can be presented in various styles, but can also be mapped.



Dataprovider has crawled over 67 million websites (by May 2014) and provides a huge amount of data/information for their clients. The company is also able to provide organised business data for professional databases, for example credit and marketing databases.

Dataprovider's general director visited the Hanze University of Applied Sciences in November 2013 for a guest lecture. Students were asked to think about a new business concept for providing data for the business-to-consumer market, a totally new market for Dataprovider.

Students of the 4th year bachelor specialisation 'Information and Innovation' wrote papers using the 'business canvas model' for working out a new business concept.

For further details see Dataprovider (<http://www.dataprovider.com>).



The screenshot shows the Dataprovider website. At the top, there is a navigation bar with links for 'Tour', 'Business info', 'SEO tools', 'TLD', 'Hosting', 'Contact', 'SIGN UP', and 'Log in'. The main heading reads 'We crawl the web and structure the data'. Below this, a sub-heading states 'Dataprovider transforms the internet into a structured information database to create insightful information about companies.' A 'LEARN MORE' button is visible. A large graphic features the word 'dataprovider' with a play button icon. At the bottom of the page, it says 'Today we have 67,299,550 websites available for your selection' and 'Search our database for prospects'.

The screenshot shows a Vimeo video player. The video content is a network diagram with a central figure of a woman thinking, surrounded by nodes and lines. A banner in the center of the diagram says 'new kind of bubble plastic'. The video player interface includes the Vimeo logo, search bar, and playback controls.

The screenshot shows a news article from the website 'versionN'. The article title is 'Danske konsulenter vil score kassen på frigivne geodata'. The author is 'EMNER Digitalisering, Offentlig it' and there are '22 kommentarer'. The article is published on '30. jul. 2013'. The logo for 'SEPTIMA' is prominently displayed. Below the article, there is a section for 'BØRSEN.' with various market-related links like 'NYHEDER', 'KURSER', 'AVISEN', etc.

EXAMPLE APPLICATION ON FREE ELEVATION DATA (FOR SPILLAGES, VIEW CHANGES)

An example for the use of Dataprovider data is a private company, who have built various solutions based on open data. Below is an example where they use the Danish elevation model to illustrate what will happen if we have a car accident, where oil is spilled on the ground.

LIQUID SPILLS

Septima Labs'Any Fluid' is an example of how Septima can make simulations in real time on the large amounts of terrain data.

See the demo: <http://labs.septima.dk/flow/index.html>

How to operate the demo?

Click "Mark the center of the spill" and then on the map somewhere in Odsherred where you want to simulate a spill.

What does the demo do?

Once the user has entered the spill center, it forms a circle around the center with a radius of 10 meters. The service then simulates how a liquid that "let loose" on this circle would flow.

What data is the service using?

The simulations are made on an improved version of the Danish Digital Elevation Model. Denmark Elevation Model is the official elevation model for Denmark and can be downloaded free from Geodatastyrelsens Site Map Supply

As Denmark's elevation model is produced with hydraulic simulations in mind, Septima made numerous improvements to the elevation model before using it for fluid simulations. What improvements needed to be carried out depended on what the simulations would be used for. In this demo, there is an emphasis on the simulation to show where the spill could potentially end up, in which river, lake or sea.

The improved elevation model in this demo is only available for Odsherred (the bright area of the card).

Other options

In this demo, a circle with a fixed radius around the selected point was used. The calculation engine can also handle the user indicating the position of the



precise extent of the spill.

The simulation in the demo shows the river or lake spill in a worst case scenario. Simulations can also be made that show the spill will be retained by depressions in the ground before it reaches the stream or lake.

There are other examples from the papers such as in the image below:

