D2.1 SOCIAL AND OPEN DATA VISUALISATION METHODS AND DATA SOURCES REPORT

PROJECT

Acronym: UrbanData2Decide
Title: Data Visualisation and Decision Making Solutions to Forecast and Manage Complex Urban Challenges
Coordinator: SYNYO GmbH

Reference: 847511
Type: Joint Programme Initiative
Programme: Urban Europe

Start: September 2014
Duration: 26 months

Website: http://www.urbandata2decide.eu
E-Mail: office@urbandata2decide.eu

Consortium: SYNYO GmbH, Research & Development Department, Austria (SYNYO)
University of Oxford, Oxford Internet Institute, UK (OXFORD)
Malmö University, Department of Urban Studies, Sweden (MU)
Open Data Institute, Research Department, UK (ODI)
IT University of Copenhagen, Software Development Group, Denmark (ITU)
ZSI Centre for Social Innovation, Department of Knowledge and Technology, Austria (ZSI)
DELIVERABLE

Number: D2.1

Title: Social and Open Data Visualisation Methods and Data Source Report

Lead partner: OXFORD

Work package: WP2: Basic Exploration, Stakeholder Studies and Requirement Analysis

Date: February 2015

Authors: Jonathan Bright, OII
Julia Neuschmid, SYNYO
Ulrich Atz, ODI

Contributors: Taha Yasseri, OII
Stefano de Sabbata, OII
Tom Heath, ODI
Ariën de Klerk, SYNYO

Reviewers: Markus Rasmusson, Lunds University
Nicklas Guldåker, Lunds University
Per Olof Hallin, Malmo

The UrbanData2Decide project is co-funded under the Joint Programming Initiative, 2nd call Urban Europe.
# TABLE OF CONTENTS

1 Introduction ........................................................................................................................................... 6  
1.1 Scope ............................................................................................................................................... 6  
1.2 Methods ......................................................................................................................................... 6  
2 Urban Decision Making – A Cycle Process ....................................................................................... 7  
3 Social Media Data for Urban Decision Making ............................................................................... 11  
3.1 Social Media Data for Urban Decision Making: Key Opportunities ............................................ 11  
3.2 Social Media Data for Urban Decision Making: Key Challenges .................................................. 15  
4 Open Data for Urban Decision Making ............................................................................................ 16  
4.1 Current status of Open Data in European cities ............................................................................. 17  
4.2 Open Data for Urban Decision Making: Key opportunities ......................................................... 25  
4.3 Open Data for Urban Decision Making: Key challenges ............................................................... 36  
5 Summary of Applications for Open Data and Social Media Data ..................................................... 38  
6 Data Visualisation Methods for Urban Decision Making .................................................................. 40  
6.1 Content Based Visualisations ........................................................................................................ 40  
6.2 Map Based Visualisations ............................................................................................................ 44  
6.3 Linked Geovisualisation ................................................................................................................ 49  
7 Summary ............................................................................................................................................ 49  
References ............................................................................................................................................. 51  
Annex ..................................................................................................................................................... 56  
Abbreviations ....................................................................................................................................... 56  
Glossary of Terms ................................................................................................................................. 56
LIST OF FIGURES

Figure 1: The cycle urban decision making process (based on Schrenk et al., 2011) ...................... 7
Figure 2: Phases of the Renewal of the Vienna Mariahilfer Straße (based on City of Vienna, 2014) .... 8
Figure 3: Vienna Shopping Street Mariahilfer Straße before (left) and after (right) the reconstruction (Sources: viennaphoto.wordpress.com, kurier.at) ......................................................................................... 9
Figure 4: General Framework of urban decision making (based on UN HABITAT 2001, Kingston et al. 2005, Schrenk et al., 2011) .................................................................................................................. 10
Figure 5: Country clusters based on Open Data Barometer Readiness and Impact Questions (The World Wide Web Foundation, 2015) ........................................................................................................ 18
Figure 6: Open data readiness in different regions (Davies, 2013) ......................................................... 19
Figure 7: Country Analysis United Kingdom 2014 (Source: ODI, 2014) .................................................. 21
Figure 8: Country Analysis Sweden 2014 (Source: ODI, 2014) ............................................................... 22
Figure 9: Country Analysis Denmark 2014 (Source: ODI, 2014) ............................................................. 22
Figure 10: Country Analysis Austria 2014 (Source: ODI, 2014) ............................................................... 23
Figure 11: Snapshot of Open Data Index for different countries ............................................................. 24
Figure 12: Open data based routing application WANN .............................................................................. 26
Figure 13: Example of a visualisation application for open geographic data (Greater London Authority, n.d.a) ........................................................................................................................................ 28
Figure 14: Example Whereabouts London interface (Catapult, n.d) .......................................................... 29
Figure 15: London School Atlas (Greater London Authority, 2014a) ....................................................... 30
Figure 16: London Open Workspaces (Greater London Authority, 2014b) ................................................. 31
Figure 17: Open Data on bicycle accidents in Berlin (Hörz, 2015) ................................................................ 33
Figure 18: Open Data on trees that might cause allergies in Vienna (Kiefer, 2013) ........................................ 34
Figure 19: Interface of the London development database (Greater London Authority, n.d.b) ................. 35
Figure 20: Collected applications per city ................................................................................................ 38
Figure 21 (left): Applications collected for different thematic areas .......................................................... 38
Figure 22 (right): Applications using different data: Open Data and Social Media Data .......................... 38
Figure 23: Overview on collected applications for open data in selected cities ........................................ 39
Figure 24: Word Cloud Created with Wordle ............................................................................................ 40
Figure 25: Stream Graph .......................................................................................................................... 41
Figure 26: Sunburst Graph ....................................................................................................................... 42
Figure 27: Pie chart (Source: metadata-census.com) .................................................................................. 43
Figure 28: Radar charts (left) and Isotype (right) (Source: modcult.org, www.opendataresearch.org) ................................................................................................................................................. 43
D2.1 VISUALISATION METHODS AND DATA SOURCES

Figure 29: Network visualisations (Source: socialwebthing.com, blogs.oii.ox.ac.uk) ........................................ 44
Figure 30: Chloropleth Graph Created with ArcGIS ................................................................. 45
Figure 31: Kernel Density Map Created with ArcGIS ................................................................. 46
Figure 32: Hexagonal Density Map Created with ArcGIS ............................................................ 47
Figure 33: Cartogram: regional geography of peer-to-peer lending in the UK (Source: http://smtm.labs.theodi.org) ........................................................................................................ 47
Figure 34: Variations of non-contiguous and Dorling maps (Source: www.gislounge.com, kelsocartography.com) ........................................................................................................... 48
Figure 35: The GeoVIZ toolkit ........................................................................................................ 49

LIST OF TABLES

Table 1: Examples of applications using social media data ............................................................. 14
Table 2: Examples of applications using open data in transport and mobility ................................. 27
Table 3: Examples of applications using open data in the field of society ..................................... 28
Table 4: Examples of applications using open data in the field of culture .................................... 29
Table 5: Examples of applications using open data in the field of education and sports ............... 31
Table 6: Examples of applications using open data in the field of economy ................................. 32
Table 7: Examples of applications using open data in the field of security and safety .................... 33
Table 8: Examples of applications using open data in the field of health ........................................ 34
Table 9: Examples of applications using open data in the field of urban planning .......................... 35
1 INTRODUCTION

1.1 Scope

This report summarises background research undertaken by the UrbanData2Decide project team into social media data and open data; research which will be used as an evidence base for work package 3 – method modelling, visualisation design and framework elaboration – concerning the design of the overall system. The report looks in particular at two main issues: data sources (sections 3 and 4) as well as its applications (section 5), and data visualisation methods (section 6). In both cases, we focus on the use case of urban decision making, hence distinguishing our report from more general literature on social media, open data and visualisation.

In terms of data sources, our interest lies in mapping out what data is available which could be integrated into the project from social media platforms and open data (especially open government data), and also exploring other uses which have been made of the data by similar projects. In terms of visualisation methods, we explore some of the most common techniques, with a particular focus on techniques for visualising text content and techniques for mapping, making use of example and toy datasets to illustrate their various strengths and weaknesses.

1.2 Methods

Data sources – Social Media data and Open (Government) data – and visualisation methods were described using desk research as well as literature research focusing on reports from the World Wide Web Foundation and Open Data Institute.

Applications from different domains that use Open Data and Social Media Data were primarily collected through Internet and literature research using platforms and open data catalogue websites of cities such as the one that is provided by the City of Vienna (n.d). Applications were collected according to a defined template specifying each tools name, a short description, the data type, the field of application, such as security and safety, environment, transport and mobility, etc., and the city where the application can be used. The aim was to show the great variety in which open data and social media data can be used. The way the applications have been evaluated and classified is described in detail in chapter 5.
2 URBAN DECISION MAKING – A CYCLE PROCESS

Urban decision making is not a single event, but can be described as a dynamic process, structured through different and progressive phases. Ideally it can be considered as a cycle process that starts with issues identification, base analysis and stakeholder identification and mobilisation, continues with data collection and analysis, strategy development and implementation which on its turn leads to monitoring from where the cycle will start all over again. The urban decision making process is not a linear one but ongoing, meaning that it gets continuous input (socio-demographic data, economic data, changes in infrastructure, environmental data, etc.) which constantly influences the stakeholders, plans, explanatory reports, etc. and decisions. Figure 1 describes the urban decision making cycle as an interaction of several stakeholders (e.g. from public administration on national, regional, local level, research, industry, policy makers, the public, media, etc.) that generate certain dynamics.

![Urban Decision Making Cycle](image)

**Figure 1: The cycle urban decision making process (based on Schrenk et al., 2011)**

It needs to be highlighted that in the real world usually not each of these phases is taken into consideration. For example in many cases there is no monitoring phase or often decisions are done ad hoc without an in-depth analysis of the current situation. But when we speak about a transparent and evidence-based decision making process, key elements are the conceptual phase that includes balancing of different interests, the decision phase, and the implementation in the real world. Urban development is not a one way street leading towards an optimal end. Changes of the framework like revival or crisis of the economic sector, the change of social ideals or ecologically driven challenges, ask for the adaption of the system, its aims, strategies and concepts. Therefore, urban decision makers rely on regular input of sound data and information. New data and information need to be
collected and analysed to be able to detect changes and to estimate the impact these changes could have in the monitoring phase. (Schrenk et al., 2011) The following example shows the renewal and stakeholder involvement process of a major street in the city of Vienna.

**Real world example:** The Mariahilfer Straße is the biggest shopping street in the City of Vienna and was heavily frequented by cars and other motorized vehicles until it was decided recently to transform most parts of the street, including its parking spaces, into a pedestrian area and a shared space area. Since then Vienna’s major shopping street Mariahilfer Straße has been undergoing the whole scale revitalization. The street is now turning from a car-busy street into a pedestrian friendly area with several advantages: broader sidewalks, new concepts for outdoor terraces and street cafes, more space for pedestrians, new consumer-free seating areas, new lightning, Wifi installation, noise reduction, traffic calming, inclusive street design, etc. Nevertheless, the reconstruction of the street faced at its beginning a lot of criticism and political controversies such as the fear of an increase in traffic and noise in the surrounding streets and loss in sales during the reconstruction phase that was even discussed on a national level.

**Figure 2: Phases of the Renewal of the Vienna Mariahilfer Straße (based on City of Vienna, 2014)**

The issue first came up in 2010/2011 after the city council elections when the Social Democratic Party started to govern in coalition with the Green Party. After first baseline studies were conducted, the stakeholder involvement process (see figure 2) started in 2011 with a number of big stakeholder involvement events including more than 600 participants (citizens) that contributed with more than 1,000 ideas for the reconstruction of the street. In addition to the physical meetings, an online dialogue with more than 900 posts was moderated. In 2012 selected experts were integrated in the stakeholder involvement process in the form of round tables. An extra survey for all residents in the relevant districts was conducted in 2013. In the same year further stakeholder involvement was done to collaboratively design the new Mariahilfer Straße with citizens, experts and urban planners. In late 2013 first pilots were implemented and tested. The feedback had influence on the final detail planning and reconstruction of the street. The first major section of the street was completed by the end of 2014. In 2015 all parts of the street are planned to be fully reconstructed. (City of Vienna, 2014)
Increasing technical opportunities open the door for new management tools to support processes of improved urban governance. A stream of guidelines and tools that support decision making processes have been developed (refer to UrbanData2Decide Deliverable D2.2). They can be used for reporting, public information or consultation, data analysis, data visualisation, monitoring, or modelling. Such tools can also support the operational principles of good urban governance as described by UN Habitat (2001). The operational principles are:

(1) Civic engagement: People are the principal wealth of cities; they are both the object and the means of sustainable development. Civic engagement implies that living together is not a passive exercise: in cities, people must actively contribute to the common good.

(2) Equity: The sharing of power leads to equity in the access to and use of resources. Women and men must participate as equals in all urban decision-making, priority-setting and resource allocation processes. Inclusive cities provide everyone – be it the poor, the young or older persons, religious or ethnic minorities or the handicapped – with equitable access to nutrition, education, employment and livelihood, health care, shelter, safe drinking water, sanitation and other basic services.

(3) Transparency: The accountability of local authorities to their citizens is a fundamental tenet of good governance. Transparency and accountability are essential to stakeholder understanding of local government and to who is benefiting from decisions and actions. Access to information is fundamental to this understanding and to good governance.

(4) Security: Every individual has the inalienable right to live, liberty and the security of person. Insecurity has a disproportionate impact in further marginalising poor communities. Cities must strive to avoid human conflicts and natural disasters by involving all stakeholders in crime- and conflict prevention and disaster preparedness.

(5) Subsidiarity: Responsibility for service provision should be allocated on the basis of the principle of subsidiarity, that is, at the closest appropriate level consistent with efficient and cost effective
delivery of services. Cities should be empowered with sufficient resources and autonomy to meet their responsibilities.

(6) Efficiency: Cities must be financially sound and cost-effective in their management of revenue sources and expenditures, the administration and delivery of services, and in the enablement, based on comparative advantage, of government, the private sector and communities to contribute formally or informally to the urban economy.

(7) Sustainability: Cities must balance the social, economic and environmental needs of present and future generations. Leaders of all sections of urban society must have a long-term, strategic vision of sustainable urban development and the ability to reconcile divergent interests for the common good.

Figure 4 summarises the general framework of urban decision making including the different phases, normative goals, stakeholders, and decision support tools, as described above. In addition, it shows that urban decision making is an interdisciplinary task and tackles different sectors of activity or thematic areas, e.g. transport and mobility, societal challenges, environment, and the economy.

Figure 4: General Framework of urban decision making (based on UN HABITAT 2001, Kingston et al. 2005, Schrenk et al., 2011)

A further discussion on urban decision making processes is done in Deliverable 2.2. To be highlighted is the strong need of sound data and information for evidence-based decision making processes from different thematic areas and sources. Data sources are mainly public administrations on local, regional and national level, research sector and industry, and more and more citizens in the form of user-generated content.
3 SOCIAL MEDIA DATA FOR URBAN DECISION MAKING

At a basic level, social media can be defined as media where those using the system are also those who create the content. It is this characteristic which distinguishes them from other types of media (such as the news media), where content consumers and content producers play different roles (though of course news outlets are increasingly incorporating many social features into their offerings). In addition, many social media sites permit users to create and maintain lists of contacts with whom they want to share content with; lists which are often based on social ties formed elsewhere; some have defined these more restrictively as social networking sites, though in this report we will continue to refer to them as social media.

Of course, under this definition, many types of media (such as email or even the telephone) are essentially “social”. However much of the current academic and policy interest around social media stems from the recent spread and mass uptake of a small number of relatively new social media websites, of which Twitter and Facebook are the primary examples. Two main factors have generated this trend. First, unlike (for example) email platforms, these social media platforms have adopted a quite open stance to sharing the data created on their service. Partly as a way of spreading usage, and partly as a way of encouraging third parties to develop applications which make use of their systems, Twitter and Facebook (and others) have made some portions of their data available through “Application Programming Interfaces” [APIs]: which has encouraged researchers to use this data as a way of better understanding users of these platforms. Second, these platforms are increasingly reaching very high penetration rates in many countries around the world. This mass penetration creates the possibility that content created on these platforms will offer increasing insight into what is going on in society as a whole (though we raise important challenges to this claim below).

Owing to both of these factors, a wide variety of projects have been launched which attempt to use social media data for a wide variety of different purposes. In what follows, we first map out the key projects and ideas which have been created of relevance to urban decision making. We then look at some of the key challenges and barriers still to be overcome.

3.1 Social Media Data for Urban Decision Making: Key Opportunities

In this section, we look at some existing work using social media data with specific relevance to the context of urban decision making, and hence highlight some key opportunities presented by this type of data source. We do this under two headings: crisis detection and management; and census and survey data.
Crisis Detection and Management

Probably the biggest relevant research area which social media has stimulated has been in the domain of crisis detection and management (Kavanaugh et al., 2012). By providing live access into the thoughts, feelings and opinions of citizens, social media are uniquely positioned to be able to highlight unusual events or crises as they emerge, or perhaps even before then. Importantly, this can take place much faster than conventional crisis highlighting mechanisms.

This application has been shown in a variety of domains. In terms of public health, studies have successfully applied Google Search data (Ginsberg et al. 2008) and Twitter data (Signorini et al., 2011) to detect influenza type outbreaks and epidemics (though recent studies have highlighted that Google Search data may be decreasingly useful in this regard – see Lazer et al. 2014). Such mechanisms are useful because they are able to report instantly; whereas flu statistics collected from local hospitals and clinics might suffer a time lag of several weeks. Similarly, other research has applied Twitter data to the case of dengue fever (Gomide et al., 2011). Early detection is especially important in this case because it permits the targeting of insecticide and bug spraying activities, which can suppress epidemics but only if they are launched quickly enough. Similarly, Sina Weibo data has also been analysed for its potential usefulness as a pollution sensor (Tse et al., 2014), by looking at the extent to which pollution related words in these social media postings relate to actual pollution outcomes.

In terms of natural disasters, real-time earthquake detection has been an especially important type of case study (see e.g. Sakaki, Okazaki and Matsuo 2010; Earle, Bowden and Guy 2011). Again, systems here offer the potential to be much quicker than conventional detection methods, particularly in terms of their ability to send targeted warnings to people potentially in the area of the earthquake. This would be especially useful in countries with high volumes of earthquakes (e.g. Japan).

Finally in this context, social media data have been applied to crime detection and prevention, particularly in the context of riots and urban disturbances. Many police forces already actively monitor social media channels during times of unrest; and researchers have furthermore shown how computers could potentially monitor such channels passively, waiting for any particular outbreak of violence (Jones and Palazzolo 2014).

In all of these applications, the methodology is broadly speaking quite similar. A signal is detected within social media postings, highlighting either a topic which people are talking about (for example, pollution) or a particular sentiment (for example, anger). The means of doing so technically range from simple keyword detection (e.g. mentions of the word “smog”) to more sophisticated machine learning techniques which rely on training the computer to classify posts as either relevant or not relevant (though, of course, techniques are still developing here). This signal is then compared to
existing data sources for validation. For example Google Flu trends data is compared to actual disease reports from the US Centre for Disease Control.

**Social and Demographic Data Capture**

A second and more developing area of research has been in the use of social media data as a substitute for traditional data capture techniques, such as censuses and surveys, which are costly and time consuming to implement. Social media data offer potentially cheaper and faster solutions, as well as offering the possibility of capturing data which traditional data collection instruments would struggle to identify.

The area of transport data is one key example of this. Official collection of transport statistics can be slow and cumbersome, based around surveying residents and asking questions about their commute. However, social media data can provide real-time insight into actual traffic patterns, in a way that does not require any outright commitment or contribution from commuters. For example, one study used a dataset of 172 million trips taken in around 14,000 New York Taxi cabs to explore how lift sharing could decrease traffic within the city (Santi et al., 2014); and we should highlight that the development of social taxicab apps such as Uber could provide even more data in that regard. Other studies have made use of Twitter for similar types of traffic monitoring purpose (Ribeiro et al., 2012; Hochmair and Cvetujevic 2014).

Researchers have also been exploring the use of social media data in order to reveal broad scale demographic information. For example, Liu, et al. (2014) are able to successfully estimate population densities in different regions of Australia using geotagged tweets. Jurdak, et al. (2014) extend this approach to looking at population mobility within Australia. Although their work takes place at the high level of aggregation, it has potential application to the city case studies highlighted in UrbanData2Decide.

Beyond population location, researchers have also looked into more specific geographical characteristics, for example the different types of uses which are made of different areas of cities (see Cranshaw et al., 2012; Rösler and Liebig 2013; Zhi et al. 2014; Frias-Martinez et al. 2014). For example, Hochman and Manovich (2013) compare the, what they call, “visual signatures” of 13 different global cities using 2.3 million Instagram photos from these cities. Meanwhile, Silva et al., (2012) make use of Foursquare to propose a technique called city image and show its applicability taking as examples eight different cities. The resulting image is a way of summarizing the city dynamics based on transition graphs, which map the movements of individuals. Similarly, Kling and Pozdnoukhov (2012) apply content analysis to Foursquare and Twitter to detect the “stories” that the city tells. And finally Grinberg et al. (2013) use the same type of data as Kling and Pozdnoukhov but with focus on diurnal patterns and types of activities based on the time of the day, the authors
detect different functionalities of different regions in the city. More specially they use visualisations of over 200,000 Instagram photos uploaded in Tel Aviv, Israel over three months to show how they can offer social, cultural and political insights about people’s activities in particular locations and time periods. Such data might be revealed through census instruments, however censuses are typically taken very rarely (e.g. every 10 years), which means they can quickly become out of date, especially in a context of high migration. Other examples in this regard include the use of social data in the forecasting of unemployment (Llorente et al., 2014), in the monitoring of forest environments (Daume et al., 2014), or even the general detection of “irregularities” related to events in the city (Lee and Kazutoshi, 2010).

Finally, in this context, it is worth mentioning the use of mobile phone data. Mobile phones are somewhat outside the remit of social media described above, as they are usually quite closed systems which do not freely share their data. However, researchers which have been able to gain access have reported important results in their ability to estimate population levels and demographic data (Deville et al., 2014; Bogomolov et al., 2014; Lukowicz et al., 2014).

While the applications here are somewhat different, in methodological terms the approaches taken are quite similar. Social media postings are collected and then a certain type of signal is extracted (for example, the graphical location of users making the postings). Then, the signal (or more exactly the distribution of this signal) is compared to actual known distributions, for example census data on the distribution of the population.

**Table 1: Examples of applications using social media data**

<table>
<thead>
<tr>
<th>City</th>
<th>Domain</th>
<th>Application</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>London</td>
<td>Culture</td>
<td>Twitter</td>
<td>Map of different languages using used in twitter in London</td>
<td><a href="http://twitter.mappinglondon.co.uk/">http://twitter.mappinglondon.co.uk/</a></td>
</tr>
<tr>
<td>Rotterdam</td>
<td>Economy</td>
<td>Jobtease</td>
<td>Online searching for jobs, get a message when you pass a company that hires people</td>
<td><a href="http://www.rotterdam.nl/mailing:jobtease">http://www.rotterdam.nl/mailing:jobtease</a></td>
</tr>
<tr>
<td>London</td>
<td>Health</td>
<td>Sickweather</td>
<td>Health forecasting on social media</td>
<td><a href="http://www.sickweather.com/">http://www.sickweather.com/</a></td>
</tr>
<tr>
<td>Vienna</td>
<td>Transport and</td>
<td>Westbahn</td>
<td>Analysis of social media content on public transport (new railway)</td>
<td><a href="http://www.webmarketingblog.at/2011/12/21/social-media-monitoring-westbahn/">http://www.webmarketingblog.at/2011/12/21/social-media-monitoring-westbahn/</a></td>
</tr>
</tbody>
</table>
### 3.2 Social Media Data for Urban Decision Making: Key Challenges

Social media data hence offers a wide variety of opportunities for policy makers and decision makers in urban environments. However the use of this data is also not without its challenges. In this section, we highlight three key ones: its demographic representativeness, its accessibility and its long term sustainability.

#### Demographic Representativeness

A key concern for both researchers and decision makers using social media is the extent to which users of social media can accurately be said to represent the population as a whole. Social media penetration levels are high, yet are also unevenly distributed throughout the population, particularly tilted towards younger age groups (OXIS 2013); whilst geotagging is also very patchy and uneven. Furthermore, not everyone who uses social media creates equal amounts of content: rather, the majority of content is created by the minority of users. All of this creates important questions about the consequences of using social media data for urban decision making. For example, using it for crisis management might exclude certain underrepresented groups from receiving government help, whilst using it for demographic statistics might paint a misleading picture.

#### Accessibility

A second problem concerns the accessibility of social media data. Many social media platforms make data available for free (though under a specific licence or service level agreement), but do not guarantee its availability over the long term. Rather, this availability is designed to encourage others to develop applications which enhance the social media platform’s overall business model. Changes to this model might result in changes to the type of data being made available. Furthermore, data availability is often quite limited: for example, Twitter makes tweets available as they are being created, but offer very limited options for long term historical search. This means those using the data need to plan in advance to capture it as it is created.
Long term sustainability

A final point, related to the above, is the long term sustainability of applications developed using social media data. As we highlighted, recent research has called into question the current usefulness of predicting flu outbreaks using Google Search, which was previously one of the standout applications of social media data (Harford 2014). A variety of reasons have been offered as to why the system appears to be breaking down: one of the most important being changes to the way Google itself interacts with its users (in particular, the appearance of a facility which suggests searches as users are typing). As the business models of social media platforms evolve, it may be that the content they produce evolves as well, changes which may come to invalidate applications which were developed to rely on them.

All three of these challenges highlight the need for social media data to be benchmarked continually: that is, validated against existing trusted sources of data and other ways of measuring the phenomenon in question (for example, by matching earthquake predictions against those recorded by seismological agencies).

4 OPEN DATA FOR URBAN DECISION MAKING

We will now move on to examine our second major data source, that is open data. Open (Government) Data [O(G)D] platforms such as by the Government of the United Kingdom (n.d), the Austrian Government (Bundeskanzleramt Österreich, n.d) or an European data portal, providing access to open, freely reusable datasets from local, regional and national public bodies across Europe (Open Knowledge Foundation, n.d) make datasets available to the public. The majority of open data comes from central government departments, a number of other public sector bodies and local authorities, but also from research (Kauppinen, 2012), industry and the non-governmental sector (NGOs). We also have to mention that the open data movement is diverse and made up of a range of constituencies with different agendas and aims, which are not driven by one party but several national, regional and local initiatives (Kitchin 2013). Data exists for all different kinds of themes, e.g. environment, health, sports and leisure, education, transportation, etc., and often with a spatial reference to be used not only in tabular form but also in form of cartographic representations. This discussion on open data shows that there are several characteristics of open data, but not one common definition.

The Open Data Institute (ODI) writes on their website: Open data is data that anyone can access, use and share.¹ This is not the only definition of open data. The GovLab has compiled an overview on

¹ http://theodi.org/guides/what-open-data

© 2015 UrbanData2Decide | Urban Europe
their blog. In summary, the definitions center on similar themes, but have different flavours when it comes to specific elements such as licensing or costs. For example, the ODI states that a license is a necessary requirement for open data.

Most proponents of open data would argue that the prize should reflect the marginal cost, which is often zero because of the digital infrastructure. However, this does not imply that publishing or using open data does not require resources. The cost may therefore be substantial, especially for first-time publishers or users, because of the setup cost, infrastructure, time and technical expertise that may be required. Moreover, the provision of data at zero cost still allows for business models modelled on a freemium approach or charging users via a service-level agreement.

4.1 Current status of Open Data in European cities

The current status of open data is described in the following major reports that are Open Data Barometer and the Open Data Index.

Open Data Barometer

There are already several reports and publications analysing the adoption of open data across the globe: The Open Data Barometer 2013 (Davies 2013) and 2015 (The World Wide Web Foundation, 2015) focus on the analysis of open government data (OGD) across the world in terms of context, availability, and emerging impacts. They provide a snapshot of OGD practices at national level and outline a country-by country ranking. Covering a broad sample of 77 countries in 2013 (Davies 2013) and 86 countries in 2015 (The World Wide Web Foundation, 2015), these reports combine peer-reviewed expert survey data and secondary indicators to look at open data readiness, implementation and emerging impacts.

Figure 5 illustrates the diversity of the OGD landscape in terms of adoption and readiness of open data and the level of activity of individual governments. Based on an evaluation of a number of models in the Open Data Barometer Report, a four-cluster analysis was selected and, based on a detailed review of qualitative and quantitative data in each cluster, labeled: (1) High-capacity; (2) Emerging & advancing; (3) Capacity constrained; and (4) One-sided initiatives.

2 http://thegovlab.org/open-data-whats-in-a-name/
Legend: **High capacity, emerging & advancing, capacity constrained, one sided initiatives**

**Figure 5: Country clusters based on Open Data Barometer Readiness and Impact Questions (The World Wide Web Foundation, 2015)**

**High-capacity** - Countries such as UK, US, Sweden, France all have established open data policies, generally with strong political backing. They have extended a culture of open data out beyond a single government department with open data practices adopted in different government agencies, and increasingly at a local government level.

**Emerging & advancing** - Countries such as Spain, Czech Republic, Italy, Portugal, Greece, Ireland have emerging or established open data programmes, often as dedicated initiatives, but sometimes through linking open data into existing policy agendas.

**Capacity constrained** - Countries such as Turkey, Ukraine in this cluster all face challenges in establishing sustainable open data initiatives as a result of limited government, civil society or private sector capacity, limits on affordable widespread Internet access, and weaknesses in digital data collection and management.

**One-sided initiatives** - These countries each have some form of open data initiative, ranging from departmental web pages listing open data, to full open data portals. However, government action to publish selected datasets is not matched by civil society capacity and freedom to engage with the data, nor by private sector involvement in the open data process. (The World Wide Web Foundation 2015)

Figure 6 uses radar charts to illustrate the readiness of OGD in different regions. From the charts, we can see that, Europe is the leading region, while the deployment of OGD in Middle East & Central Asia and Africa is very limited (Davies 2013).
A global movement to make government “open by default” picked up steam in 2013, when the G8 leaders signed an Open Data Charter (Cabinet Office, 2013) — promising to make public sector data openly available, without charge and in re-useable formats. In 2014 the UN recognised the need for a “Data Revolution” to achieve global development goals (UN, 2014). However, this second edition of the Open Data Barometer shows that there is still a long way to go to put the power of data in the hands of citizens. Core data on how governments are spending our money and how public services are performing remain inaccessible or pay walled in most countries. Information critical to fight corruption and promote fair competition, such as company registers, public sector contracts, and land titles, is even harder to get. In most countries, proactive disclosure of government data is not mandated in law or policy as part of a wider right to information, and privacy protections are weak or uncertain (The World Wide Web Foundation 2015). The Open Data Barometer further identifies that (Davies, 2013; The World Wide Web Foundation, 2015):
• OGD policies have seen rapid diffusion over the last years, reaching over 55% of the countries surveyed in the Barometer 2013. The OGD initiatives launched have taken a range of different forms: from isolated open data portals launched within an e-government framework to ambitious government-wide OGD implementations.

• Global progress towards embedding open data policies stalled in 2014. While many countries with moderate or strong OGD initiatives in 2013 have seen steady growth in the availability and impacts of OGD, a number of countries have slipped backwards over the last 12 months. Many of the countries that made initial steps with OGD in 2012/13 have not sustained their OGD commitments and activities. Government that is “open by default” is a long way off for most of the world’s citizens.

• Much more needs to be done to support data-enabled democracy around the world. There has been very limited expansion of transparency and accountability impacts from OGD in 2014.

• To maximise impact, open data needs go local. Political impacts from open data are greater in countries that have city-level open data activities.

• For data to be considered truly open, it must be published in bulk, machine-readable formats, and under an open license. In 2015, just over 10% of the 1,290 different datasets surveyed for the Barometer met these criteria — a small but significant increase from 2013, when 7% of datasets were published in full open data format. Thirty-one countries have at least one open dataset, and just over 50% of the datasets surveyed among the 11 top-ranked countries qualified as fully open.

• Leading countries in the ODB are investing in the creation of ‘National Data Infrastructures’ to provide a foundation for public and private innovation and efficiency. Examples are UK, Sweden, USA, Canada, Australia, Norway, Denmark, Germany, France, etc. They have high-level and broad-based political backing for the OGD initiatives, and are investing in capacity building with entrepreneurs and intermediaries. They are also focusing on building communities around open data, convening government officials and outside stakeholders to understand more clearly how data can be harnessed for economic and social progress. However, no countries can yet claim to fully be ‘open by default’, and embedding OGD practices across government is a key future challenge.

• Mid-ranking countries such as Switzerland, Italy, Austria, Ireland, etc. have put in place some of the components of an OGD initiative, such as an open data portal and competitions or events to catalyse re-use of data, but have often failed to make key datasets available, and are lacking in important foundations for effective open data re-use. Absence of strong Right to Information laws may prevent citizens from using open data to hold government to account, and weak or absent Data Protection Laws may undermine citizen confidence in OGD initiatives. In addition, limited training and support for intermediaries might result in data which cannot be mobilised to generate economic and social benefits. Strong regional differences have to be taken into account.
• Low-ranking countries have not yet started to engage with Open Data, and many developing countries lack basic foundations such as well-managed and digitised government datasets. In these countries, interventions to support OGD may look radically different from the leading OGD initiatives surveyed in the Barometer – with opportunities for open data approaches to generate, as well as use, public information.

• The Barometer 2015 ranks the UK as the most advanced country for open data readiness, implementation and impact, scoring above the USA (2nd), New Zealand (3rd), Sweden (4th), The Netherlands and France. Just 16 of the 77 countries (20%) included in the 2013 Open Data Barometer saw a reduction in their scaled ODB score in this 2014 edition. In general, the trend is towards steady, but not outstanding, growth in open data readiness and implementation.

The UK, USA and Sweden remain at the top of the Barometer overall. Each country has placed an emphasis on the economic growth potential of open data and, over the last year, each has continued to develop mechanisms for engaging with private sector data users — from the Open Data User Group in the UK, to the Open Data Forum convened by the Ministry of Enterprise and Innovation in Sweden, and the Open Data Roundtables series convened by the GovLab at NYU in partnership with the US Federal Government. Denmark is ranked 9th being strong in the readiness of the data as well as economic and political impacts, Austria is ranked 15th from 86 countries, being strong in social impacts but weak in economic impacts. The detailed country analysis of the UrbanData2Decide partner countries (UK, Sweden, Denmark and Austria) are provided below (figures 7-10):

Figure 7: Country Analysis United Kingdom 2014 (Source: ODI, 2014)
Denmark lost 4 ranks, and is now ranked number 9. Denmark experienced modest reductions in its scores and rankings — mostly as a result of weaker implementation — which appear to be in part correcting for some over-scoring of dataset openness in this country in 2013.

Austria’s rank changed from 18 to 15 on the Barometer ranking between 2013 and 2015. After the federal election in late 2013, Austria’s new government included open data in its coalition agreement (Austrian Federal Chancellery, 2003), but researchers reported that, as of August 2014, no member of the cabinet could be identified as in charge of the subject. In general, the Austrian open data
agenda appears to be driven by several major cities and regions; in centres such as Vienna, start-up activity around open data is generating social, economic and environmental returns.

Figure 10: Country Analysis Austria 2014 (Source: ODI, 2014)

Open Data Index

The Open Data Index was launched in 2014 by the Open Knowledge Foundation as a mechanism to assess the state of open data around the world. It covers information concerning the data sets published by national governments in over 110 countries (70 in 2013). Annual snapshots of the data are presented on the Website to showcase the results of the project. One of the main goals of this project is to stimulate debate and action between citizens and their governments to lead to the release of further data assets. From the 1668 key datasets which have been identified in the current release of the index, only 222 data sets are considered open (13%).

The Open Data Index structures the open data sets into 10 different categories such as transport timetables, budget, election, national map, etc., and each category uses the same nine criteria to measure the availability of the data, such as whether the data is online and free of charge. Based on the submissions provided by the editors, the index gives a total score of openness for each country. Figure 11 demonstrates a snapshot of the countries with top scores in Open Data Index. From the snapshot we can see that five European countries are in the top ten of the index. Furthermore, the Open Data Index also provides a detailed breakdown score for each country in each category.
Potential of Open Data

In addition to the previous reports there are a number of reports from consultant or market research institutes such as Deloitte and McKinsey, emphasising on the potential of Open Data. Deloitte’s ‘Open Data Driving Growth Ingenuity and Innovation’ looked at the open data landscape as of 2012 to identify trends for the future development of this area and recommendations for the commercial sector. They estimate that in the near future business will engage in open data in four aspects: (1) strategically exploit the rapidly growing of their open data assets; (2) opening up their data assets as a revolution way of competing; (3) using open data to improve transparency and engage customers; and (4) work with government and make policies for data responsibility and privacy (Deloitte, 2012).

A report produced by McKinsey sought to “quantify the potential value of open data by examining applications in seven sectors of the global economy”: education, transportation, consumer products, electricity, oil and gas, healthcare, and consumer finance (Manyika, 2013). The report encompasses not only an analysis of the economic value of open data in each of these sectors, but also a discussion

---

**Figure 11: Snapshot of Open Data Index for different countries**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 United Kingdom</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>970</td>
</tr>
<tr>
<td>2 Denmark</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>825</td>
</tr>
<tr>
<td>3 Finland</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>730</td>
</tr>
<tr>
<td>4 Australia</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>720</td>
</tr>
<tr>
<td>5 New Zealand</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>720</td>
</tr>
<tr>
<td>6 Norway</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>710</td>
</tr>
<tr>
<td>7 United States</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>700</td>
</tr>
<tr>
<td>8 Germany</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>690</td>
</tr>
<tr>
<td>9 India</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>675</td>
</tr>
<tr>
<td>10 Taiwan</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>670</td>
</tr>
<tr>
<td>11 Colombia</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>660</td>
</tr>
<tr>
<td>12 Czech Republic</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>655</td>
</tr>
<tr>
<td>13 Sweden</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>655</td>
</tr>
<tr>
<td>14 Uruguay</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>655</td>
</tr>
<tr>
<td>15 Romania</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>640</td>
</tr>
</tbody>
</table>

© 2015 UrbanData2Decide | Urban Europe
of potential barriers to adoption and actions to be taken to ensure that this potential is not lost. For example, the privacy issues are major concerns in nearly all the investigated domains and appropriate legal and regulatory frameworks are urgently needed to ensure that open data is distributed in an anonymous and secure manner. In the Perini (2013) it is highlighted that there is a need for methodological and tool support to allow the “various stakeholders to engage in an informed dialogue and to guide the future development of open data”.

The Open Government Data Stakeholder Survey 2010 (Martin et al., 2011) led by the LOD2 project focused on open government data. They report on the requirements of different stakeholder groups (citizens, public administration agencies, policy makers, industry, media, and science) regarding open data sets and catalogues. From the survey, national and regional data sets are most required by the stakeholders and they demand more data to be published in non-proprietary formats such as CSV and XML.

4.2 Open Data for Urban Decision Making: Key opportunities

In this section we look at some existing work using open data with specific relevance to the context of urban decision making, and hence highlight some key opportunities presented by this type of data source.

Citizen participation and collaboration

The open data-trend has been picked up by a number of cities and – together with networks, software and innovation-friendly legal standards – can be considered as part of the fundament of innovative city governance (European Innovation Partnership, 2015). As regards the definition of open data, however, it is crucial to note that in spite of its widespread use, there is no common understanding as to the type of data (raw data versus processed or aggregated data), topical foci (e.g. traffic, mobility, health data, etc.) or the target group of this data (citizens generally, businesses, web-developers, NGOs, etc.). Due to this broad definition, numerous urban initiatives labelled “open data” can be found, which differ in all of these aspects but share the characteristics of an information source that is meant to serve the broader public. In the governmental context, this indicates that parts of government data shall be made freely accessible to citizens (Kuhn, 2014) – an aspect that is often discussed under the label “open government” and “government as a platform” (O’Reilly, 2011). Yet, the concepts of open data and open government are not interchangeable, but in general, the first is the precondition for the latter: Citizens shall be given the possibility to contribute to government initiatives and to interact with the public sector, which requires sufficient information, e.g. through open data (e.g. Chan, 2013). For pointing out this interrelation, we are using the term “open government data” for open data published by public agencies or governments.
participation and collaboration of citizens is a possible new paradigm of governance that can be enabled by open government data.

The general way of thought is that open (government) data could enable forms of collaborative and participative governance since, in order for citizens to participate in public projects or to voice their opinion, they first need to learn about the addressed issue and also need to have a platform where they can share their contributions. From this perspective open data can serve both to lower the barriers for participation and collaboration and to make citizen involvement more attractive (Jetzek et al., 2013). Open government data might enable new and more participative and collaborative forms of governance as it has been voiced by many scholars in the field (Ferro et al., 2013; Jetzek et al., 2013; O’Reilly, 2011). Bartenberger and Grubmüller-Régent (2014) propose that there are especially three elements where collaborative governance can benefit from open government data: overcoming knowledge asymmetries, facilitating joint fact finding and enabling trust building. First, the usage of open government data in collaborative governance could decrease knowledge asymmetries and thereby level the playing field. This could mean for example that a city administration releases datasets that are relevant for a certain forum of collaborative governance on their open data portal. Second, open government data could support joint fact finding by providing shared knowledge bases and accepted sources of information that can be used and accessed by all stakeholders. Third, open government data could be an additional factor that promotes mutual trust and understanding especially on a symbolic level. By making parts of their knowledge available to all stakeholders and the public as open data, state actors can show that they have nothing to hide and are willing to commit to the process of collaboration. (Bartenberger and Grubmüller-Régent, 2014)

**Transport and mobility**

The most popular use of many cities open data has been the creation of mobile journey planning apps. Especially with the release of real-time data live applications become possible. Developers have created hundreds of applications, reaching millions of active users in several European cities. Popular applications are routing and location-based services such as public transportation real-time monitors answering questions such as “Where is the next public transport station? When will the next bus/tram arrive? etc.”. Further there are public transportation journey planners to get from A to B, journey planners for bikes, applications for efficient car parking and payment management (short-term parking zones), or applications that help to find the next taxi. For example WANN is a live public transportation service available in several European cities such as Vienna, Linz, Geneva and Berlin using real-time and open data.

**Figure 12: Open data based routing application WANN**
### Table 2: Examples of applications using open data in transport and mobility

<table>
<thead>
<tr>
<th>City</th>
<th>Application</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin</td>
<td>wheelmap.org</td>
<td>Open map to search, find and mark accessible places for wheelchair users</td>
<td><a href="http://wheelmap.org/">http://wheelmap.org/</a></td>
</tr>
<tr>
<td>London</td>
<td>London JamCams</td>
<td>Shows all traffic jams through live feed</td>
<td><a href="http://data.gov.uk/apps/london-jamcams-iphone-app">http://data.gov.uk/apps/london-jamcams-iphone-app</a></td>
</tr>
<tr>
<td>London</td>
<td>Roadwork Database</td>
<td>Find all current road works around you and plan your trip carefully</td>
<td><a href="http://data.gov.uk/apps/roadworks-database">http://data.gov.uk/apps/roadworks-database</a></td>
</tr>
<tr>
<td>Manchester</td>
<td>DataGM</td>
<td>Public transport data for Manchester</td>
<td><a href="http://datagm.org.uk/">http://datagm.org.uk/</a></td>
</tr>
<tr>
<td>Manchester</td>
<td>PayByPhone</td>
<td>Pay a parking fee through the app</td>
<td><a href="http://www.manchester.gov.uk/info/471/parking_in_public_areas/5897/pay_by_phone">http://www.manchester.gov.uk/info/471/parking_in_public_areas/5897/pay_by_phone</a></td>
</tr>
<tr>
<td>Rotterdam</td>
<td>SuperB</td>
<td>Shows in real-time boats on the Maas river. In this game, the user is the captain</td>
<td><a href="http://www.rotterdam.opendata.org/es/web/guest/app;jsessionid=223BE819894197CA53102AAABB18F61/?state=getAppSlider">http://www.rotterdam.opendata.org/es/web/guest/app;jsessionid=223BE819894197CA53102AAABB18F61/?state=getAppSlider</a></td>
</tr>
<tr>
<td>Vienna</td>
<td>Wann Today</td>
<td>Public transportation real-time monitor (next public transport station, next bus/tram, etc.)</td>
<td><a href="http://subzero.eu/wann/?lang=en">http://subzero.eu/wann/?lang=en</a></td>
</tr>
<tr>
<td>Vienna</td>
<td>ParkSheriff</td>
<td>Car parking and paying (short-term parking zone)</td>
<td><a href="https://open.wien.gv.at/site/parksheriff-handyparken/">https://open.wien.gv.at/site/parksheriff-handyparken/</a></td>
</tr>
</tbody>
</table>

### Society

The population development of many European cities and regions is quite dynamic, whereas some face population growth, others are confronted with population decline. With this comes a range of challenges to tackle, e.g. meeting the need for affordable housing, providing the right amount of school places for children, and ensuring capacities of transport networks. Nowadays, decision makers can look at open demographic statistics and population projections with interactive tools to find out about the expected population growth in the local areas and can therefore base decisions, drive and steer large-scale developments such as in the Barking Riverside or Olympic Park legacy developments in London (Greater London Authority, n.d.a).
Figure 13: Example of a visualisation application for open geographic data (Greater London Authority, n.d.a)

Table 3: Examples of applications using open data in the field of society

<table>
<thead>
<tr>
<th>City</th>
<th>Application</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin</td>
<td>Berlinwahlkarte</td>
<td>Elections (visualisation of results)</td>
<td><a href="http://www.morgenpost.de/berlinwahlkarte">http://www.morgenpost.de/berlinwahlkarte</a></td>
</tr>
<tr>
<td>London</td>
<td>Flatmate Rooms</td>
<td>Find shared apartments or flatmates</td>
<td><a href="http://data.gov.uk/apps/flatmate-rooms-easy-roommate-and-spare-room-finder-uk-wide">http://data.gov.uk/apps/flatmate-rooms-easy-roommate-and-spare-room-finder-uk-wide</a></td>
</tr>
<tr>
<td>London</td>
<td>London Data Store Visualiser</td>
<td>Visualisation and analysis of changing characteristics of the population</td>
<td><a href="http://data.london.gov.uk/case-studies/population-projections/">http://data.london.gov.uk/case-studies/population-projections/</a></td>
</tr>
<tr>
<td>Rotterdam</td>
<td>Fixmymap</td>
<td>Play against your neighbours and report as many defects as possible, such as broken garbage cans or graffiti paintings, all reports go directly to the municipality</td>
<td><a href="http://www.rotterdamopendata.org/es/web/guest/app;sessionid=2238E819894197CA53102AA398B18F61/?state=getAppSlider">http://www.rotterdamopendata.org/es/web/guest/app;sessionid=2238E819894197CA53102AA398B18F61/?state=getAppSlider</a></td>
</tr>
</tbody>
</table>
Culture

Culture can be an opportunity for a city to generate considerable educational, social and economic benefits and it can help foster urban regeneration, change the city's image and raise its visibility and profile on an international scale. A great number of applications have been developed using open data about museums, memorials and other cultural sites in cities, often targeted at tourists and integrated in mobile tourist guides. Another example on how open data can be used is Whereabouts London (Catapult, n.d) that is an ongoing experiment to understand and improve future cities. By blending 235 types of data, the project is investigating what London could look like if we drew its boundaries afresh, grouping neighbourhoods based on how people live – not where they live. Results are eight clusters in London based on similarities between people and places. With reimagining neighbourhoods local authorities could work out how to share their services with each other, transport providers could tailor their services to travellers better than ever, behavioural change campaigns could be targeted in new ways to make them work more effectively. Whereabouts London uses data from the Greater London Authority’s Datastore that is a hub for data about all aspects of the city. Users can find, explore and build on over 500 different datasets that the city generates, by either downloading them or accessing them through APIs.

Figure 14: Example Whereabouts London interface (Catapult, n.d)

Table 4: Examples of applications using open data in the field of culture

<table>
<thead>
<tr>
<th>City</th>
<th>Application</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>London</td>
<td>Whereabouts</td>
<td>Societal clusters</td>
<td><a href="http://whereaboutslondon.org/#/map">http://whereaboutslondon.org/#/map</a></td>
</tr>
<tr>
<td>Rotterdam</td>
<td>Doek voor in je hoek</td>
<td>Analyses your interior and gives you 5 paintings (from the Rijksmuseum) that would suit</td>
<td><a href="http://www.rotterdamopendata.org/es/web/guest/app.jsessionid=223BE819894197CA53102AAA8B18F61/?state=getAppSlider">http://www.rotterdamopendata.org/es/web/guest/app.jsessionid=223BE819894197CA53102AAA8B18F61/?state=getAppSlider</a></td>
</tr>
</tbody>
</table>
### Education and sports

The pattern of which schools serve an urban area is extremely complicated, but with the School Atlas application, for the first time, educational planners and parents are able to see the picture across London. The open London Schools Atlas projections show London may need up to 4,000 extra primary classes across the capital by the end of the decade. In addition to catchments, the map shows school performance data and population projects for school age children (Greater London Authority, 2014a).

![London School Atlas](https://example.com/london-school-atlas.png)

*Figure 15: London School Atlas (Greater London Authority, 2014a)*

<table>
<thead>
<tr>
<th>City</th>
<th>Application Name</th>
<th>Description</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stockholm</td>
<td>Historiska Stockholmsbilder</td>
<td>Find out what was located at your current location in the past, see historic pictures</td>
<td><a href="https://www.slowtravelstockholm.com/2014/10/20/stockhols-best-apps/">https://www.slowtravelstockholm.com/2014/10/20/stockhols-best-apps/</a></td>
</tr>
</tbody>
</table>
Table 5: Examples of applications using open data in the field of education and sports

<table>
<thead>
<tr>
<th>City</th>
<th>Application</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin</td>
<td>Kindergarten-Suche</td>
<td>Search for kindergarten</td>
<td><a href="http://www.tursics.de/kindergarten/de/">http://www.tursics.de/kindergarten/de/</a></td>
</tr>
<tr>
<td>London</td>
<td>School Atlas</td>
<td>Map out patterns of demand of schools in an interactive map</td>
<td><a href="http://data.london.gov.uk/case-studies/school-atlas/">http://data.london.gov.uk/case-studies/school-atlas/</a></td>
</tr>
<tr>
<td>London</td>
<td>Intelligent London</td>
<td>Analyse and visualize the skills of young Londoners</td>
<td><a href="http://data.gov.uk/apps/intelligent-london">http://data.gov.uk/apps/intelligent-london</a></td>
</tr>
<tr>
<td>Vienna</td>
<td>Büchereien in Wien</td>
<td>Library finder</td>
<td><a href="https://open.wien.gv.at/site/buechereien-in-wien/">https://open.wien.gv.at/site/buechereien-in-wien/</a></td>
</tr>
<tr>
<td>Manchester</td>
<td>NH connect</td>
<td>Find a league or facility where you can play your favorite sport</td>
<td><a href="http://manchesterinklink.com/city-launches-">http://manchesterinklink.com/city-launches-</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>manchester-nh-connect-mobile-app/</td>
</tr>
</tbody>
</table>

**Economy**

New businesses require working spaces. *London Open Workspaces* is an open interactive map of more than 130 start-up friendly office spaces and 200 artist studios to help small and medium sized enterprises (SMEs) find the right type of space.

![Image of London Open Workspaces](image_url)

**Figure 16:** London Open Workspaces (Greater London Authority, 2014b)
Users can search by factors including price, length of stay and types of business already in the building. Providers of workspaces are encouraged to keep this a live and accessible database. This data has been published to help SMEs themselves but will also be used by the London Enterprise Partnership to inform future investment. By making the information publically available, as well as the services they provide, the London Open Workspaces map is an invaluable resource for start-ups, SMEs, artists, and many more. (Greater London Authority, 2014b)

**Table 6: Examples of applications using open data in the field of economy**

<table>
<thead>
<tr>
<th>City</th>
<th>Application</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin</td>
<td>Marktverzeichnis Berlin &amp; Brandenburg</td>
<td>Mapping of markets in Berlin and attributes (opening times, etc.)</td>
<td><a href="http://markt-verzeichnis-bb.de">http://markt-verzeichnis-bb.de</a></td>
</tr>
<tr>
<td>London</td>
<td>Where Can I Live?</td>
<td>Tells you what place you could afford and what is close to your job</td>
<td><a href="http://data.gov.uk/apps/where-can-i-live">http://data.gov.uk/apps/where-can-i-live</a></td>
</tr>
<tr>
<td>London</td>
<td>London Open Workspaces</td>
<td>Helps SMEs find the right type of space (start-up friendly office spaces and artist studios)</td>
<td><a href="http://data.london.gov.uk/case-studies/workspaces/">http://data.london.gov.uk/case-studies/workspaces/</a></td>
</tr>
<tr>
<td>Stockholm</td>
<td>Börstkankes appen</td>
<td>Give away and receive used items in an easy way</td>
<td><a href="http://www.openstockholmaward.se">http://www.openstockholmaward.se</a></td>
</tr>
<tr>
<td>Vienna</td>
<td>Interaktive Budgetvisualisierung</td>
<td>Interactive Budget Visualisation using Google Motion Charts</td>
<td><a href="https://open.wien.gv.at/site/interaktive-bud">https://open.wien.gv.at/site/interaktive-bud</a> getvisualisierung-mit-google-motion-charts/</td>
</tr>
</tbody>
</table>

**Safety and security**

With the opening of data on crimes and accidents several maps have been developed for different cities showing the city administration, urban planners, the police and other emergency organisations where resources should be concentrated, while the public can identify risky areas to avoid and demand more police action if necessary. An example in the field of safety and security is an open data application from Berlin that shows bicycle accidents in the city on a map. It is accessible for the public as well as for the city administration and transportation planners who can use the visualised data to define new strategies and improved safety measures for individual transportation (cars, bikes, pedestrians). (Hörz, 2015)
Figure 17: Open Data on bicycle accidents in Berlin (Hörz, 2015)

Table 7: Examples of applications using open data in the field of security and safety

<table>
<thead>
<tr>
<th>City</th>
<th>Application</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin</td>
<td>Berliner Fahrradunfälle</td>
<td>Visualisation of bicycle accidents</td>
<td><a href="http://daten.berlin.de/anwendungen/berliner-fahrradunf%C3%A4lle">http://daten.berlin.de/anwendungen/berliner-fahrradunf%C3%A4lle</a></td>
</tr>
<tr>
<td>Manchester</td>
<td>NH connect</td>
<td>Report local issues</td>
<td><a href="http://manchesterinklink.com/city-launches-manchester-nh-connect-mobile-app/">http://manchesterinklink.com/city-launches-manchester-nh-connect-mobile-app/</a></td>
</tr>
<tr>
<td>Stockholm</td>
<td>Resledaren</td>
<td>Helps people with a cognitive handicap to get from point A to B safely</td>
<td><a href="http://www.openstockholmaward.se/">http://www.openstockholmaward.se/</a></td>
</tr>
<tr>
<td>Vienna</td>
<td>Fundboxen in Wien</td>
<td>Hand in found items</td>
<td><a href="https://open.wien.gv.at/site/fundboxen-in-wien/">https://open.wien.gv.at/site/fundboxen-in-wien/</a></td>
</tr>
</tbody>
</table>

Health

A healthy city is one that is continually creating and improving those physical and social environments and expanding those community resources which enable people to mutually support each other in performing all the functions of life and developing to their maximum potential (WHO, 1998). With opening health data several applications to show and support health of citizens have been developed. The open data application Pollenradar Wien (Kiefer, 2013) shows trees that can cause allergies in Vienna and is enriched with several attributes such as blooming period.
application brings valuable information for citizens who plan trips, sports, and other outdoor leisure activities, and who suffer from pollen. The app uses open data from the city of Vienna.

Figure 18: Open Data on trees that might cause allergies in Vienna (Kiefer, 2013)

### Table 8: Examples of applications using open data in the field of health

<table>
<thead>
<tr>
<th>City</th>
<th>Application</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin</td>
<td>Ozon Sonar</td>
<td>Visualisation of ozone data</td>
<td><a href="http://ozon.sonar1.mobi/berlin/">http://ozon.sonar1.mobi/berlin/</a></td>
</tr>
<tr>
<td>London</td>
<td>AirText</td>
<td>Information about air quality for people who suffer from asthma, emphysema, bronchitis, heart disease or angina</td>
<td><a href="http://data.gov.uk/apps/airtext">http://data.gov.uk/apps/airtext</a></td>
</tr>
<tr>
<td>Manchester</td>
<td>Light Raider</td>
<td>Motivation app, 'collect' street lamps while jogging, try to collect as much as possible to beat other joggers.</td>
<td><a href="http://www.manchestereveningnews.co.uk/business/business-news/10-mobile-apps-created-greater-8123152">http://www.manchestereveningnews.co.uk/business/business-news/10-mobile-apps-created-greater-8123152</a></td>
</tr>
<tr>
<td>Stockholm</td>
<td>Cykeland</td>
<td>Triggers people to use the bike, game based and real-time info about routes etc.</td>
<td><a href="http://www.openstockholmaward.se/">http://www.openstockholmaward.se/</a></td>
</tr>
<tr>
<td>Vienna</td>
<td>Familiengesundheit</td>
<td>Health planner (find POIs, diary, health checklist, etc.)</td>
<td><a href="https://open.wien.gv.at/site/familiengesundheit">https://open.wien.gv.at/site/familiengesundheit</a></td>
</tr>
</tbody>
</table>

### Urban planning

Mobile applications are useful tools to engage citizen in urban planning processes. The city of Berlin uses the open data application *buergerbautstadt* that on the one hand informs citizens about planned urban developments and constructions. On the other hand citizens can provide their input and feedback to these projects as part of stakeholder involvement and citizen participation processes. (Noffke, n.d)

Another example of an open data application in urban planning comes from London. While most of the London Boroughs provide details of the planning applications they receive on their own websites, it can be difficult to get an overview across London or to view the permissions granted in a particular location where it is close to borough boundaries. The London Development Database (LDD) records selected planning permissions in the Greater London Area (GLA) as part of the process of monitoring the Mayor’s London Plan. The data is made available to the public and professionals in the
development field via the “London Development Database Webmap” website, in which they can search by area or postcode to find the location of live and completed planning permissions anywhere in Greater London. (Greater London Authority, n.d.b)

![Image of the London Development Database Webmap](image.png)

**Figure 19: Interface of the London development database (Greater London Authority, n.d.b)**

<table>
<thead>
<tr>
<th>City</th>
<th>Application</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin</td>
<td>BürgerBautStadt</td>
<td>App for civic engagement in urban planning processes</td>
<td><a href="http://buergerbautstadt.de">http://buergerbautstadt.de</a></td>
</tr>
<tr>
<td>London</td>
<td>London Development Database</td>
<td>Visualisation of planning permissions, residential units, floor space, open spaces</td>
<td><a href="https://www.london.gov.uk/priorities/planning/london-development-database">https://www.london.gov.uk/priorities/planning/london-development-database</a></td>
</tr>
<tr>
<td>London</td>
<td>London Rents Map</td>
<td>Visualisation of average private sector rents for different types of home</td>
<td><a href="http://data.london.gov.uk/case-studies/rents-map/">http://data.london.gov.uk/case-studies/rents-map/</a></td>
</tr>
<tr>
<td>Rotterdam</td>
<td>Omgevingsalert App</td>
<td>Alerts you when something is happening in your surroundings from a tree that is being cut to new urban projects</td>
<td><a href="http://www.rotterdamopendata.org/es/web/guest/app?sessionid=223BEB19B94197CA531DZZAABBB18F611?state=getAppSl">http://www.rotterdamopendata.org/es/web/guest/app?sessionid=223BEB19B94197CA531DZZAABBB18F611?state=getAppSl</a> ider</td>
</tr>
<tr>
<td>Vienna</td>
<td>Accessible Vienna</td>
<td>Accessibility of buildings</td>
<td><a href="https://open.wien.gv.at/site/accessible-vienna/">https://open.wien.gv.at/site/accessible-vienna/</a></td>
</tr>
</tbody>
</table>
4.3 Open Data for Urban Decision Making: Key challenges

Recent initiatives have dramatically increased the range of previously “closed” data being made “open” by the government, including data sets on travel, weather and healthcare. Without any doubt, access to data allows people to work together more effectively, collaborating with each other, with policy-makers and with service providers to improve governance, public life and public services to make more informed decisions in cities. However, the open data initiative also raises some concerns and criticism. Challenges to face when dealing with open data include the following:

Data needs to be actionable

There is a clear and compelling case that information produced at public expense should be made open and freely available to benefit the public. However, simply declaring data sets to be open in itself does not make it of any practical use to the public. When released in its raw form, data is not open to the public in any meaningful sense. It is only open to a small elite of technical specialists who know how to interpret and use it, as well as to those that can afford to employ them. Therefore, data needs to be made easy-to-use (or actionable) and public awareness and training needs to take place to enable communities to apply data to solve local problems. According to Roberts (2012) Governments should be required to release data in actionable formats conforming to open data standards - and to be fair there is already progress in this regard. But comparatively little is being done at community level to promote the re-use of public data for public benefit. Almost nothing is being done to create capacity within communities to interpret and apply open data themselves, without creating technical dependencies. To maximise the public benefit derived from public data Roberts (2012) highlights we must raise community awareness about the potentials of open data and develop the practical skills and capacities so that those potentials are realised in practice. Data is all very well, but the ability to extract meaningful information from it requires considerable skills; without it, incorrect conclusions may be drawn (Hand 2012; Kitchin 2013).

Costs

Collecting, 'cleaning', managing and disseminating data are typically labour- and/or cost-intensive processes. Open data might well be a free resource for end-users, but its production and curation is certainly not without significant cost (especially with respect to appropriate technologies and skilled staffing). In many cases, the data being opened has to date been a major source of revenue for organisations, and in the case of companies, competitive advantage. A key question, therefore, centres on how open data projects are funded sustainably in the absence of a direct revenue stream? (Kitchin 2013) A number of different models have been suggested (Ferro and Osella 2013), and it is acknowledged that securing a stable financial base is best achieved by direct government subvention.
Feedback effects on data quality

The very act of publishing the data can influence the quality of future data. Crime maps illustrate the sort of problems that can arise: The creation of crime maps have been a clear public benefit, showing on the one hand the police, emergency organisations, private security organisations, urban planners, public administration, and other stakeholders in the field of safety and security where resources should be concentrated. Visualising crime data on a map combined with other data (e.g. population, land use, housing, transportation network, public buildings, schools, etc.) helps stakeholders to improve their plans, strategies and tactics. On the other hand the public can identify risky areas to avoid and demand more police action if necessary. Nevertheless, this can also lead to the stigmatisation or downgrading of certain neighbourhoods. A survey found that respondents claim to have seen but not reported an incident in their street because they feared it would make it more difficult to rent or sell their house (Hand 2012).

Great potentials when treated with care

Much more critical attention needs to be paid to how open data projects are developing as complex socio-technical systems with diverse stakeholders and agendas. To date, efforts have concentrated on the political and technical work of establishing open data projects, and not enough on studying these discursive and material moves and their consequences. As a result, we lack detailed case studies of open data projects in action, the assemblages surrounding and shaping them, and the messy, contingent and relational ways in which they unfold (Kitchin, 2013). At the same time, an argument for open data is not only that public money was used to fund the work and so it should be universally available, but there are many potential gains and values from Open Data Initiatives: This data can be used by anyone to create great new products, business opportunities and community services. Open data enables accountability: it is difficult to conceal something if the facts are there for all to see. Open data empowers communities: crime rates, educational achievement, social services and so on are laid bare. Open data drives economic growth: more small companies are using open data to build innovative applications. They are creating new forms of transparency and accountability, fostering new form of social participation and evidence-informed modes of governance, and promoting innovation and wealth generation. Open data may even lead to more accurate conclusions and better decisions, as a wider variety of interested parties have the opportunity to examine the facts. No technology is without concomitant risks, but provided we tread carefully, with an awareness of the problems, the open data initiative holds immense promise for a better society. (Hand, 2012; Huijboom and Van der Broek, 2011; Janssen, 2012; Kitchin, 2013; Yiu, 2012)
5 SUMMARY OF APPLICATIONS FOR OPEN DATA AND SOCIAL MEDIA DATA

A number of 143 applications were systematically collected to show how open data and social media data can be used in different domains. Thereby we focused on the local level on different European cities, i.e. Vienna, London, Berlin, Manchester, Rotterdam, Stockholm, and on the European level on European wide applications (Figure 20). Selected examples for social media data applications are presented in chapters 3.1 and for open data applications in chapter 4.2 of this report.

![Figure 20: Collected applications per city](image)

![Figure 21 (left): Applications collected for different thematic areas](image)

![Figure 22 (right): Applications using different data: Open Data and Social Media Data](image)
The applications where collected for different thematic fields that are transport and mobility, society, culture, education, sports, environment, economy, security and safety, health, urban planning and other. 21% of the collected applications are in the field of transport and mobility (figure 21). Most popular applications are routing and location-based services such as public transportation real time monitors and multi-modal journey planning apps for bikes, cars and public transportation making use of real-time data.

Most of the collected applications make use of Open Data (91%) whereas only 9% use Social Media Data such as Facebook and Twitter data (see figure 22). Figure 23 presents a summary of the data collection, and highlights where applications that make use of open data and social media data could be found in each city.

![Figure 23: Overview on collected applications for open data and social media data in selected cities](image)

In terms of open data we can clearly see that there already exists a huge amount of open data applications in every city (green points). In addition to open data, many of these applications make use of other data sources, e.g. user generated data. Whereas a large amount of applications for open data can be found easily, applications that use social media data for urban topics are still hard to find (blue points). We can see differences between cities, London being one of the cities where most applications for social media data could be collected. In the fields of education, environment and urban planning no applications that make use of social media data could have been collected at all. Anyway, applications which use social media data could be identified in the fields of transport and mobility, society, sports, economy, security and safety, and also culture as well as health.
6 DATA VISUALISATION METHODS FOR URBAN DECISION MAKING

In this part of the report, we move on to describing data visualisation methods for urban decision making. In it, we look at several common techniques, providing examples and discussing the advantages and disadvantages of each. The visualisations displayed were all created using an example dataset of geo-located tweets drawn from within the city of Vienna on November 16th, 2014. However, they are broadly applicable to a wide variety of different contexts.

We divide our techniques into two broad sections: content based visualisations, which summarise and describe text created on different platforms, and geographical visualisations, which describe the geographical location of content being created.

6.1 Content Based Visualisations

In this section, we look at three common ways of visualising or summarising large amounts of text: through word clouds, stream graphs and sunburst graphs.

Word Clouds

Word clouds, such as the example one shown in the figure below, visualise the frequency of different keywords found in a particular piece of content or dataset (in this case a set of geotagged tweets).

Figure 24: Word Cloud Created with Wordle
Their principle advantage lies in the ease with which they can be understood. They may also be useful for identifying the most common words in a given dataset, which perhaps could be selected through an interactive display. However they also have a number of disadvantages: presented this way, words are taken out of context, without any semantic understanding. This issue can be lessened by using spatialization techniques (Skupin and Fabrikant 2003), to group semantically related words, and display them accordingly. It is also difficult to estimate the exact size of words (though a hover animation or mouse-over pop up could overcome this). Finally, it is difficult to show temporal development in this context, though interactive animations might partially overcome this limitation.

**Stream graphs**

![Stream Graph](image)

*Figure 25: Stream Graph*

A second option for visualising content is the stream graph, as shown in figure 25. This visualises the relative frequency of values over time, resolving one of the limitations of the word cloud. It is both visually attractive and very useful for identifying both peaks and troughs in a given keyword: in the graph, we can see clearly when the hashtag Syria starts to develop in importance. However like the word cloud there is still the problem of words being taken out of context and difficulties of estimating the exact size of the words in question, though again this second issue could potentially be resolved with mouse pop ups. There is also a clear limit to the amount of stacked areas which can be inserted if the graph is to remain intelligible (whereas the word cloud, comparatively, can display a lot more information). As a static diagram, a staked area chart may work better.
Sunburst Graphs

A further option we want to highlight in this regard is the sunburst graph. This is a nested pie/doughnut chart, which shows percentages of content in a hierarchical fashion, e.g. frequency of hashtags in social media content as percentage of overall hashtags, grouped by topic (though it should be noted that this graph only really works in an interactive environment). Here we can see how the overall category of “events” can be broken into several sub categories, such as “Konzert”.

These graphs can be used with either a manual or automatic ontology to classify hashtags, and provide a categorized overview of content which both contains a lot of data yet is relatively simple for the user to understand. Interaction can be used to provide further details. However this graph also has many of the limitations of the word cloud: words themselves are decontextualised, and over time development is limited. The number of categories and subcategories is also limited. Finally, this graph needs extra work to develop an ontology which works for the context in question.
Charts

The shape of the pie chart is sometimes a **doughnut** or **semi-circle** as shown in the example (fig. 27).

![Pie chart](image)

**Figure 27: Pie chart (Source: metadata-census.com)**

Isotype, in a simple application, are bar charts, where the bar has been replaced by symbols, e.g. pictograms for people (figure 28 right). Radar charts (figure 28 left) are especially useful if the data is cyclical such as weekdays. Other cases have to be considered carefully because marking comparisons is often hard, they become easily cluttered, and alternatives such as a bar chart may be superior. This chart from the Open Data Barometer works well because the levels are not overlapping.

![Radar charts and Isotype](image)

**Figure 28: Radar charts (left) and Isotype (right) (Source: modcult.org, www.opendataresearch.org)**
Network visualisations

Network visualisations present a particular type of content based visualisation: one that displays links between different nodes or actors in a dataset. Network visualisations can be both informative and revealing in this respect. However they also need to be created with care for their main message to stand out. One particularly problematic aspect is that the position of dots in a network diagram does not (necessarily) correspond with the actual position of actors.

![Network visualisations](socialwebthing.com, blogs.oii.ox.ac.uk)

6.2 Map Based Visualisations

In this second section, we move on to explore geographic ways of visualising data, looking in particular at several types of map: a chloropleth map, kernel density estimation map, hexagonal (point density) map, and cartograms.

**Chloropleth Map**

Chloropleth maps, such as that shown in figure 30, show the density or occurrence of content over geographical space, with respect to previously defined area units - most commonly, administrative areas. The map below shows which areas are producing the most tweets, scaled by the amount of inhabitants. These maps provide simple and clear visualisations, which also avoid problems for simple scatter plots which are created when multiple pieces of content are created in the same location and then visualised on top of each other.
Such maps are also ideal for comparison with other types of data which take place at the same statistical area (e.g. census data), though of course areas must not overlap. However they also have some problems: readers may perceive content as being evenly distributed throughout an area (when in fact it may well not be), and may also give very different results depending on the type of area selected (this is commonly referred to as the “modifiable areal unit problem” – see Gehlke and Biehl, 1934).

**Kernel Density Map**

A second option is the kernel density estimation map. This map shows the density of content over space, estimating the centre/kernel of activity, e.g., density of tweets. It avoids some of the problems highlighted above, avoiding in particular any artefacts created by the selection of administrative area whilst nevertheless also avoiding problems of occlusion present on simple scatter plot maps. However kernel estimation can also create misleading patterns. In particular, occurrences of the analysed phenomenon create halos that cover area where the phenomenon doesn’t occur, and this halo may be perceived as larger where more occurrences concentrate in a single location; e.g. high density near a river might result in an estimation of content within the river’s basin. Kernels are also much more difficult to use in comparison with explanatory variables.
Finally, there is the option of a point density map. This map is essentially a type of two dimensional histogram, with geolocated content aggregated into small predefined areas, with the colour of the bin reflecting the amount of content. This minimises some of the problems of the kernel density map (especially the ever expanding kernels) and also permits a quite precise location of content generation. However, again, the size of the selected bins is an arbitrary choice, and hence can create some artefacts and cannot easily be linked to explanatory variables. In particular, the map presented in this section is a hexagonal density map, which is an alternate method to the more classic square fishnet approach. The main advantage of a hexagonal map is that distances between the centre of a cell and the centre of all adjacent cells is constant, making for easy comparison. Hexagonal maps still have several disadvantages, depending on the use, including difficulties in creating a simple hierarchical structure for multi-scale analysis.
Another putative solution is the use of cartograms\(^3\): For example, in a project lead by the Open Data Institute, the regional geography of peer-to-peer lending in the UK was analysed and visualised. The data is available as open data. The cartogram scales the regions of the UK according to their relative peer-to-peer activity. London is hence larger and Scotland smaller than usual.

\(^3\) More than one flavour of cartogram exists, e.g. Dorling cartograms. http://www.ncgia.ucsb.edu/projects/Cartogram_Central/cartogram_examples/dorling3.jpg
There are many different types of cartogram, but each one works on the same basic principal of distorting a geographical map in order to convey further information about different places. A non-contiguous cartogram is the simplest and easiest type of cartogram to make. In a non-contiguous cartogram, the geographic objects do not have to maintain connectivity with their adjacent objects. This connectivity is called topology. By freeing the objects from their adjacent objects, they can grow or shrink in size and still maintain their shape. A Dorling cartogram maintains neither shape, topology nor object centroids, though it has proven to be a very effective cartogram method. To create a Dorling cartogram, instead of enlarging or shrinking the objects themselves, the cartographer will replace the objects with a uniform shape, usually a circle, of the appropriate size.  

Dorling and Dorling-like cartograms

Figure 34: Variations of non-contiguous and Dorling maps (Source: www.gislounge.com, kelsocartography.com)

---

4 See http://www.ncgia.ucsb.edu/projects/Cartogram_Central/types.html

© 2015 UrbanData2Decide | Urban Europe
6.3 Linked Geovisualisation

As we have described above, all visualisation tools have strengths and weaknesses, but none of them is perfect. In order to take advantage of the different strengths of the different visualisations, a linked visualisation interactive tool can be created. In such a system, multiple visualisations are presented in the same tool/application. When the user interacts with one of the visualisations, the other visualisations adapt. For instance, when selecting a hashtag from the wordle, a map can be modified to show only tweets containing that hashtag, and a bar chart, area chart, or stream graph can show the frequency of that particular hashtag over time.

Figure 35: The GeoVIZ toolkit

An example of this is provided by the GeoVIZ toolkit (Hardisty and Robinson 2011), as highlighted above in figure 35.

7 SUMMARY

This report has tackled three main topics. First, we have looked at the use of social media data for urban decision making, highlighting current applications of social media data (particularly in terms of crisis detection and management, and social and demographic data capture). So far social media data for urban decision making has mainly been used by academics and researchers and hardly by city governments, so that we can say that there is still potential space for future exploitation of the data.
in urban governance. We also highlighted key challenges in the areas of demographic representativeness, accessibility and long term sustainability of data, in particular the need for constant benchmarking of social media data against other slower but more reliable instruments such as opinion polls and sample surveys. The resulting application which comes out of UrbanData2Decide will need to take into account these points when deciding how to integrate social media data.

We then looked at open data for urban decision making, reviewing the current status of open data in major European cities, and noting widespread availability in the project cities of interest. We explored open data applications in a variety of areas such as citizen participation and transport and mobility. We also highlighted clear challenges: the need for data to be actionable, the costs to the city itself or preparing and releasing the data, and potential feedback effects on data quality (such as crime mapping leading to diminished crime reporting). Again, these challenges need to be taken into account when the UrbanData2Decide application itself is developed.

Both, open data and social media data are rich sources for urban decision making with great potential. The aim is not to replace traditional data collection methods such as surveys, but additionally to consult new rich sources. The UrbanData2Decide decision support tool will demonstrate the use of various sources, namely (open government) statistical data and social media data in combination with expert advisory to support holistic urban decision making processes and will test it on real world use cases.

Finally, we explored data visualisation methods for urban decision making, looking in particular and content and map based visualisations. No visualisation is perfect: rather, each one adapts better to different situations (and we also highlighted the great potential of interactive visualisations for displaying more data). This deliverable provides a menu of visualisation options from which the resulting UrbanData2Decide application could choose. All in all, this report highlights the great potential the UrbanData2Decide application can have for its future development, but also a number of important challenges it faces along the way.
REFERENCES


D2.1 Visualization Methods and Data Sources

ACM SIGSPATIAL International Workshop on Location-Based Social Networks (LBSN ’12). ACM, New York, NY, USA, 5-11.


UN, 2014. A world that counts: mobilising the data revolution for sustainable development.


ANNEX

Abbreviations

API    Application Program Interface
GIS    Geographic Information Systems
OGD    Open Government Data
SME    Small or Medium Enterprise

Glossary of Terms

Data: A value or set of values representing a specific concept or concepts. Data become “information” when analysed and possibly combined with other data in order to extract meaning, and to provide context. The meaning of data can vary depending on its context. Data includes all data. It includes, but is not limited to, 1) geospatial data 2) unstructured data, 3) structured data, etc.

Dataset: A dataset is an organized collection of data. The most basic representation of a dataset is data elements presented in tabular form. Each column represents a particular variable. Each row corresponds to a given value of that column’s variable. A dataset may also present information in a variety of non-tabular formats, such as an extended mark-up language (XML) file, a geospatial data file, or an image file, etc.

In Linked Data, a dataset means collection of RDF data, comprising one or more RDF graphs that is published, maintained, or aggregated by a single provider. In SPARQL, an RDF Dataset represents a collection of RDF graphs over which a query may be performed.

API: An Application Programming Interface (API) is an abstraction implemented in software that defines how others should make use of a software package such as a library or other reusable program. APIs are used to provide developers access to data and functionality from a given system.

5 http://www.data.gov/glossary
6 http://www.data.gov/glossary
7 http://www.w3.org/TR/ld-glossary/#dataset-rdf
8 http://www.w3.org/TR/ld-glossary/#api
Web API: An API that is designed to work over Internet\(^9\).

Open Data: A piece of data is open if anyone is free to use, reuse, and redistribute it - subject only, at most, to the requirement to attribute and/or share-alike\(^{10}\).

Open Data Resource: the datasets, their metadata and other documents published following the open data definition.

Open Government Data: Open data produced by the government. This is generally accepted to be data gathered during the course of business as usual activities which do not identify individuals or breach commercial sensitivity\(^{11}\).

Stakeholder: A person with an interest or concern in something, especially a business. In open data, a stakeholder is anybody who can affect or is affected by the publishing and consuming of open data and their indirect economic and social influences. Five generic categories of open data stakeholder categories are derived deductively:

\(^9\) http://opendatahandbook.org/en/glossary.html#term-web-api

\(^{10}\) http://opendefinition.org/

\(^{11}\) http://opendatahandbook.org/en/glossary.html#term-open-government-data