# centreforcities

# Open data or closed doors?

## Supporting research in cities

Maire Williams
December 2013

"Centre for Cities is a research and policy institute, dedicated to improving the economic success of UK cities.
We are a charity that works with cities, business and Whitehall to develop and implement policy that supports the performance of urban economies. We do this through impartial research and knowledge exchange."
**www.centreforcities.org**

CATAPULT
**Future Cities**

Supported by Future Cities Catapult

## Executive summary

**Cities are vital to the UK economy, with the 64 largest cities generating 60 per cent of GVA.**[1] They're also diverse, complex and fast-changing which creates challenges to understanding and improving them. One of the factors constraining our ability to improve the functioning of UK cities is the lack of robust and timely data at the city level.

**Fortunately two new ways of increasing the amount and improving the quality of data available on cities are emerging.** The Government's Open Data Policy, which aims to make more government data publicly and freely available to use and re-distribute; and the innovative use of big data, which covers all data collected – from CCTV cameras and sensors to mobile phones and search engine activity. While there is currently no model for estimating the size of the benefits from data re-use,[2] the 2012 Open Data White Paper noted that an Open Data Policy is associated with monetary savings, innovation and social and economic growth.[3]

**Despite the increasing importance of this data to both city and national policymaking, currently policy-makers are not fully utilising either of these methods, meaning they lag behind policy-makers in many other countries.** The UK Government needs to do more to increase awareness of and use of the data it holds. Cities themselves need to recognise the value of the data they collect via organizations such as universities and private companies and begin sharing this data for the benefit of their residents. Cities also need to be more prepared to invest in big data projects, which can generate savings in the long-run through improving existing systems without the need to invest in any major re-developments.

This report will demonstrate the necessity of high quality, timely data to help improve city economies and the lives of UK citizens. It highlights some of the benefits seen around the world from opening up data, whether that data be classed primarily as public, private or big data. It also highlights data currently held by the UK Government that if released would significantly improve the understanding of cities; and it sets out some important areas in which data collection needs to be improved.

### Policy recommendations:

- Central and local government should prioritise the release of data they hold on city economies in order to provide cities with the information needed to enhance their economic performance and meet the needs of their citizens.

- Central government and cities should set up a series of 'hack days' to encourage new uses of the data that they are releasing and secure funding from sponsors to turn ideas into functional products.

- Cities should set up special interest groups to explore what existing public datasets can be used and combined to provide new insights, savings or proto-types reflective of their city's needs.

- Cities should look to share the big data they already collect by setting up on-line platforms that allow organisations to upload their own data and use the data of others.

- Cities should lead on collecting new data and identifying uses of big data that could lead to long-term savings and benefits.

*"The focus of this report is the necessity of high quality timely data to improve understanding of cities and to develop efficient policies."*

---

1 Centre for Cities (2013) Cities Outlook – based on Primary Urban Areas
2 BIS (15 May 2013) Shakespeare review: an independent review of Public Sector Information
3 Cabinet Office (28 June 2012) Open Data White Paper: Unleashing the Potential

## Introduction

Cities are vital to the UK economy, with the 64 largest cities accounting for 58 per cent of jobs, 60 per cent of GVA, and 72 per cent of high skilled workers.[4] This means that polices that stimulate urban growth will stimulate growth of the national economy.

Cities are also diverse, complex and fast-changing, so understanding them in order to improve them is a challenge. Cities face a range of issues which limit their success and how well they function, including congestion, quality of housing and increasing pressure on services and resources due to population growth. To craft solutions to these challenges, we first need to better measure and understand them.

**Currently, the amount of robust, publicly available data at a city level is limited, both in its range and its timeliness**. For example, the data publicly available on house prices is a year old and business export data is not publicly available at a Local Authority level, despite this information being held by the Government. This places unnecessary restrictions on understanding the challenges and opportunities faced by individual cities and the ability to develop efficient policies that improve the way they function.

**Two new ways of improving the amount and quality of information available on cities can help address this problem. The first is the Government's Open Data Policy. The second is the innovative use of big data.**

Through its **Open Data Policy**, the Government has committed itself to making more of its data freely available to use and re-distribute. This commitment was made with a view to becoming one of the most open and transparent governments in the world. The 2012 Open Data White Paper noted this would generate savings, promote innovation and support social and economic growth.[5] To date over 10,000 datasets are listed as available for download via www.data.gov.uk, ranging from British road lengths by road class to TB in cattle. But despite BIS reporting that the most popular dataset category requests are location-specific data,[6] the amount of data publicly available below the regional level remains limited.

To date there has been a weakness on the demand side of the open data agenda, with the agenda being driven by what the Government wants rather than on what users want. This creates a problem because the impact of the policy is dependent on the level of demand for the datasets released. If the Government simply releases datasets that are used by very few people – either because the public is unaware the data is available or because there is limited need for the data released – by definition the benefits of an Open Data Policy will be limited.

Part of boosting demand for what's available is providing a clear, searchable list of what data is collected and what data is currently publicly available. And the UK Government has committed to improving lists of what data each department collects and is also currently consulting via the data.gov.uk site as to which datasets to prioritise releasing.

*"Currently, the amount of robust data publicly available at a city level is limited, both in its range and its timeliness"*

---

4 Centre for Cities (2013) Cities Outlook – based on Primary Urban Areas
5 Centre for Cities (2013) Cities Outlook – based on Primary Urban Areas
6 BIS (May 2013) Market Assessment of Public Sector Information.

But to ensure the full benefits of following an Open Data Policy are realised it needs to do more to raise awareness of the data that is already publicly available and the data it holds that could be made available.

**Big data** offers an alternative way to understand the challenges facing cities through use of privately held data such as mobile phone records and data collected from sensors. Cities around the world are making increasing use of this type of big data, including cities in countries such as the Ivory Coast, where location data from mobile phone records was used to determine how to reduce bus journey times through reducing congestion.

Whilst the Government's open data agenda is beginning to focus more on releasing data for which there is a known demand, big data projects have taken a more speculative approach. In these projects cities and organisations work together to collect a range of data, interlink it, then determine how this data could best be used to improve city performance. The UK currently lags behind other countries in the development and adoption of such methods, especially when compared with the United States. The progress made in other countries provides a good opportunity for UK cities to identify ways to improve the way they function.

**This report looks at the city level benefits seen around the world from the opening up of public and private data, in order to illustrate how the sharing of such data is critical for improving the functioning of our cites and life of citizens.** It reviews the Government's Open Data Strategy, the uses of big data and the main areas in which the UK's data collection needs to be improved to allow for a deeper understanding of the key issues cities face.

Section 1 highlights the benefits that sharing government data creates and the importance of engaging with organisations, institutions and citizens to increase awareness and use of the data the Government holds. It also suggests a campaign for the release of data currently held by the Government that would have a significant impact on our understanding of cities and outlines areas in which data collection needs to be improved.

Section 2 looks at the city-level benefits from the use and sharing of privately held big data. It highlights examples of countries that have seen city level improvements from either making big data fully publicly available or available at a city level.

Section 3 looks at the main barriers to overcome when sharing data, particularly privacy, as well as the cost of organising and publishing data, the lack of demand for data, the uncertainty over continuing supply of data and the lack of the necessary skills to work with big data.

Section 4 presents conclusions, emphasising that while open data policies and use of big data have already led to improvements at a city level in the UK and internationally, much more can be done, both by the Government and cities, to promote and realise the full range of opportunities these methods offer cities.

*"The sharing of public and private data is important for improving the functioning of our cites and the lives of citizens"*

## Section 1: Government data

### Better data can improve policy and promote economic performance

Various studies have reported a link between the provision and re-use of public sector information and economic growth.[7] These studies have found that improving access has a positive impact on the number of users and is associated with greater innovation, improved and new services, and efficiency improvements from better resource allocation and informed policy.

As such, there are large benefits that could be achieved from opening up city-level datasets currently held by the Government.[8] In 2012 the UK Government committed to releasing more of the data it already holds, recognising that this will improve transparency, provide citizens with the necessary data to innovate and generate savings as well as supporting social and economic growth. To date over 10,000 datasets are listed as available for download via data.gov.uk. However, despite BIS reporting that the most popular dataset category requested is location-specific data,[9] the amount of data publicly available below a regional level remains limited.

### Cities hold a wealth of data that could be used to improve their performance

London leads the way in utilising this local level data in the UK. Transport for London (TfL) is using Oyster card data to model the London transport network. This means TfL can predict the impact of delays and then prioritise investment in station upgrades, helping to ensure money is spent where it is most needed (case study 1).

*"There are large benefits that could be achieved from opening up city level datasets currently held by the government"*

**Case study 1: London, UK**

TfL currently provides information on real-time service updates and arrival and departure times of buses, train and the underground via the TfL website and phone apps. These allow travellers to plan and adapt their journeys in response to real-time transport conditions throughout London, as well as improving public scrutiny of performance.

TfL has also collaborated with UCL to make use of the vast amount of data recorded through the Oyster card system. The data has been used to monitor journey times, numbers of passengers and the effects of various situations on congestion patterns, such as the impact of mechanical failures on a particular line. This has enabled predictions to be made of how delays will impact along the network and what measures could be put in place to best limit disruptions. The data has also been used to identify where to prioritise investment in station upgrades and the impact any upgrades are likely to have. For example, TfL has estimated that, when completed, the Northern line upgrade will increase capacity on the line by 20 per cent.[10]

7 Vickery, G (2011) Review of recent studies on PSI re-use and related market developments.
8 City level refers to data provision at a Local Authority level or below.
9 BIS (May 2013) Market Assessment of Public Sector Information.
10 Transport for London (2011) Business Plan 2011/12 - 2014/15

Lambeth Council is making the data it holds publicly available through 'Lambeth in Numbers'.[11] This website provides local level data to improve understanding about how Lambeth's services can improve the wellbeing of citizens (case study 2).

### Case study 2: Lambeth, London UK

Lambeth Council has pledged to publish all the information it holds to enhance the understanding of the relationship between local assets and infrastructure and the activity and wellbeing of its population. It is developing an open data demonstrator project titled 'Lambeth in Numbers' that allows users to link up data sets about the local area. Currently, the site contains data relevant to Lambeth's Food Strategy, including DCLG's Indices of Multiple Deprivation, NHS data on child obesity, and data from Lambeth Council on food-related facilities and activity. Users are able to provide feedback on the site and can also upload their own data.[12]

## Many other countries and cities also operate an open data policy

*"Many other countries are operating an open data policy"*

An Open Data Strategy for Europe was launched by the European Commission in 2011,[13] and in June 2013 the G8 committed to moving towards open data by default.[14] Reviewing the amount of government data publicly available in other countries, as well as how it is used, suggests ways in which the UK could further improve.

At the city level, New York's Open Data Policy illustrates how the publication of government data allows organisations to link data sets together to identify previously unknown relationships. Through linking datasets from 19 different agencies on all 900,000 property lots in the city, a model was developed to predict which properties are most likely to have illegal conversions. This has allowed inspectors to target properties to visit, saving the city time and money. As illegal conversions are a high fire risk in the city, the detection of conversions has also improved housing standards and safety for residents (case study 3).

### Case Study 3: New York, USA

In 2012, New York City Council passed a law to make all City managed data available through a single web portal on or before December 31 2018 (unless that data is deemed confidential). The aim is to improve transparency and allow the public to develop innovative uses of the data. Currently data is available on a variety of themes, including public safety, education, buildings complaints, restaurant inspections and real-time traffic numbers.

The benefits seen from this include the development of a model to predict which properties are most likely to have illegal conversions, which are associated with overcrowding and fires. The Open Data policy of New York meant that data on all 900,000 property lots in the city collected from 19 different agencies could be linked together. The datasets used included information on properties where owners had unpaid/overdue property taxes, missed utilities payments and service cuts. A correlation was found between these variables and illegal conversions, allowing inspectors to target which properties to inspect. Major violations are now found at about 70 per cent of the buildings searched, rather than 13 per cent previously, saving the city time, money and improving housing standards and safety.[15]

11 http://www.lambeth-in-numbers.co.uk/
12 http://blog.boilerhouse.co.uk/2012/09/21/boilerhouse-launches-new-generation-webtools/
13 http://europa.eu/rapid/press-release_IP-11-1524_en.htm
14 Cabinet Office (June 2013) G8 Open Data Charter
15 http://www.csmonitor.com/Innovation/Tech/2013/0607/How-big-data-helps-big-cities

## The UK Government needs to go further on its open data agenda

In 2013 the Shakespeare Review noted that, to progress its open data agenda the Government should identify and release "core reference datasets".[16] In response the Government is now doing this, and is consulting via the www.data.gov.uk site as to which datasets to prioritise for release. Once identified, the aim is to release all core departmental data by 2015. As such, the Government is adopting a more demand-led release strategy to address the imbalance between supply of and demand for government-held data.

## The Government holds a number of datasets that if made publicly available could have a significant impact on understanding and improving cities.

While there are potential costs associated with releasing any data (as outlined in section 3) the data sets described in table 1 would provide more timely and detailed information than is currently available which, in turn, would improve understanding of how cities function and what changes are needed to improve city life and performance. Table 1 sets out the Government department that currently collects and holds this data at a level appropriate for city analysis (Local Authority level or below) and the benefits that would be seen from prioritising its release.

## The economic performance of UK cities could also be enhanced through increased collection, aggregation and publication of data

The Shakespeare Review noted that currently the United States invests a great deal more than the UK in its National Data Strategy and has seen large benefits from doing so in the form of world-leading business applications.[17] Comparisons with US government metro data and the work done by the Brookings Metro Monitor, which provides on-line access to a variety of data on the 100 largest USA metropolitan areas, reveal a number of key areas in which the UK could improve its provision of city data and thereby its ability to support economic growth. These are outlined in table 2.

Greater provision of this type of data would remove some of the barriers of trying to understand and evaluate city performance and would allow for targeted policy development to improve the economic performance of cities and so the UK economy.

*"There are a number of data sets held by the government that could have a significant impact on city level research"*

16 BIS (15 May 2013) Shakespeare review: an independent review of Public Sector Information
17 BIS (May 2013) Market Assessment of Public Sector Information.

*"There are a number of areas in which the UK's knowledge of cities would benefit from increased collection, aggregation and publication of data"*

## Table 1: Data to prioritise for release

| Data Source | Benefit |
|---|---|
| **Pay as you Earn (PAYE) data (HMRC)**<br><br>Data held by HMRC could give a much richer picture of the performance of, and barriers within, city labour markets. This data contains information on **wages, inequality, commuting patterns and the movement of workers** within and between sectors and cities. | • This would allow individuals to see what they could earn working in the same profession in different cities and employers could use the information when setting wages.<br><br>• Linking this data with Individual Learner Records would allow policy makers to track people moving from education/training into employment and their wage progression, so allowing them to evaluate policy interventions.<br><br>• More timely and accurate knowledge of commuting patterns would better inform policymakers of the requirements for transport and housing provision. |
| **Commercial premises data at a local authority level (held by Local Authority Planning Departments)**<br><br>There is no central database of commercial property in cities, despite it being very important to the performance of local economies. This data shows the location and supply of commercial space by type. | • This data would allow a better understanding of the workings of local property markets.<br><br>• It would allow policymakers to develop policy to address issues that these markets face, such as the over or undersupply of office space.<br><br>• Linking this to vacancy data would also allow local authorities to inform businesses of available space within their economies. |
| **Planning applications by type and size of development at a local authority level (held by CLG)**<br><br>The data provides a measure for the demand for different types of development at the city level. It would also provide a better measurement of the performance of planning departments. | • Better scrutiny of planning departments would allow national policymakers to incentivise development and penalise planning departments that did not meet national guidelines.<br><br>• This would be of benefit to both businesses and residents, as greater certainty would help encourage business investment, while any resulting investment would create a greater number of employment opportunities for residents. |
| **Export data at a local authority level (held by HMRC)**<br><br>No data is currently publicly available on exports at a city level despite it being a focus of policy in recent years. This data would show how much a city exports. | • This data would provide policy makers with a basis from which to formulate policy to support exporting, including direct export support and decisions over trade missions. |
| **Exact boundary lines of greenbelt areas (held by CLG)**<br><br>The greenbelt has a big impact on the expansion of UK cities, but the lack of freely available information about the exact location of the greenbelt within cities hinders detailed policy discussion about it. | • This data would improve the understanding of the location of existing restrictions on building around England's urban areas with implications for policymakers, analysts and developers.<br><br>• A more informed discussion around the greenbelt would allow policymakers to consider how to allow our most successful cities to expand, opening up economic opportunities that exist within them to a wider number of people, better linking businesses to workers and so supporting future economic growth. |

**Table 2: Data requests**

| Data Source | Benefit |
|---|---|
| Aggregation and publication of housing inspection scores | • This data would improve policymakers' understanding of the physical condition of the housing stock and future housing needs, allowing for detailed planning of what needs to be done to eliminate housing shortages and overcrowding and move families into quality housing. |
| Aggregation and publication of price and number of childcare facilities by area, including the number of free spaces. | • This data would make it easier for working parents to arrange childcare provision and would allow policymakers to assess the barriers parents may face in returning to work in different cities. |
| Aggregation and publication of access to and take up of employment benefits (healthcare, pensions, paid holiday and paid sick leave). | • This data would provide a measure of the rewards/ attractiveness of employment beyond wages and allow researchers to investigate the impact these have on workers' productivity.<br><br>• It would also highlight to workers what additional benefits may be available to them and may prompt those currently choosing not to work to look for work through identifying the extra gains beyond wages they might have access to if employed. |
| Collection and publication of employers' demand for skills and qualifications (wider coverage and more detailed that that currently covered in the employers skill survey). | • This data would improve the understanding of skills mis-matches/shortages by city, helping to guide the skills strategies of local authorities and Local Enterprise Partnerships. |

*"The government needs to develop the demand for open data along with its supply."*

## Supply is not enough – the Government also needs to boost demand

There are clear benefits to having an Open Data Policy. But the Government must do more than just release data; it also needs to encourage its use. The scale of benefits from an Open Data Policy are dependent on the level of demand for the released datasets. If the Government releases datasets which are used by very few people – either because they are unaware the data is available or because there is limited need for the specific data released – the benefits of the Open Data Policy will be limited. As such, the Government needs to boost demand for more data as well as supply it.

## There has been a relatively low level of demand for open data in the UK to date

To improve this, the Government is consulting as to which datasets to prioritise for release. The downside to this approach is that the potential benefit of data is not always obvious until the full dataset can be examined. This approach also assumes a detailed knowledge of the data the Government holds which is not always true. A solution to this that has proved sucessful in the US is 'Hack Days', discussed in case study 4.

Hack Days illustrate how the demand for data can be increased through engaging with individuals and companies to develop apps that present data in user-friendly ways. This not only raises awareness of what data is available, it can result in highly innovative uses of the data and can save governments money as they don't have to organize the data themselves. The resulting apps also allow subsequent users to immediately re-use the data, without first having to find the resources to link/ organize the data themselves.

### Case study 4: Hack Days

Several cities, such as Chicago and Boston have started running 'Hack Day' events and competitions, in which individuals and companies are challenged to find innovative uses of public data. This has resulted in the creation of various apps that have improved the transparency and presentation of the data, making it easy for citizens to understand and re-use the data in research. For example, The Knight News Challenge in the USA received over 800 project proposals for uses of government data - 14 of which were then provided with funding to allow their creators to produce fully functioning apps.[18] Projects included the Chicago Crash Browser. This uses the city's Transport Department data on pedestrians or cyclists injured by motor vehicles to determine where the city should invest in infrastructure upgrades to eliminate traffic accidents.

*"Big bata has been used by cities around the world to deliver more efficient and responsive services with fewer resources"*

Hack Days have been run in the UK but they have had limited success. Typically, they generate a short-term boost to open data interest and use, but because these ideas are not funded or promoted more widely they don't get developed.[19] To get the most out of Hack Days funding needs to be available from either public or private sponsors to ensure programmes that could have significant positive impacts are taken forward. Cities themselves could set up local interest groups to explore existing public datasets to discover how these could be used to provide new insights, savings or proto-types reflective of their city's needs.

To ensure the full benefits of following an Open Data Policy are realised, it is vital the Government engages with organisations such as think tanks, academic research institutions, IT and engineering companies not only to gauge the existing demand for data but to promote awareness of data that has already been released as well as data currently held but not yet released. Improved awareness and use of existing data will also help create opportunities to link current data, extend current data collections and expose duplications and inaccuracies in current datasets, leading to improvements in the efficiency of data gathering in the future.

---

18 https://www.newschallenge.org/
19 Centre for Cities interview

## Section 2: Big Data

### Issues facing cities could also be addressed using big data

Big data is defined both by its volume and its timeliness, with vast amounts of data generated every second. The term covers data collected from a wide range of sources including CCTV cameras, sensors, mobile phones and search engine activity.

Cities generate a lot of big data. However, it is typically not made publicly available or shared between organizations within the city. As such, UK cities are missing an opportunity to improve their performance.

### Big data has been used by cities around the world to deliver more efficient and responsive services with fewer resources

To date, the majority of projects have concentrated on solving transport issues, from congestion to road maintenance issues. Dublin, for example, is working with IBM to identify the cause of and solution to transport delays through sharing data collected from bus timetables, inductive-loop traffic detectors, closed-circuit television cameras and GPS updates from the city's buses. This data will be used to reduce congestion and so pollution in the city (case study 5).

> *"UK cities are missing out on an opportunity to improve their own functioning"*

**Case study 5: Dublin, Ireland**
In 2010, in order to increase the efficiency of its public transport network without the need to invest in any major re-development, Dublin city council began sharing data generated by its city services with IBM. Data is collected from bus timetables, inductive-loop traffic detectors, closed-circuit television cameras and GPS updates from each of the city's buses to identify the cause of delays and the most effective measures to put into place to improve traffic flow. From linking these datasets together, a digital map of the city has been built up and overlaid with the real-time positions of Dublin's buses. This allows traffic controllers to see the current status of the entire bus network at a glance. The data is being used to identify where additional bus lanes and bus-only traffic systems would be beneficial and if bus line start times are optimal. Meteorological data is being added to the model to allow operators to analyse the effect on the transport system of extreme weather conditions and to work to reduce the impact this has on delays. This will benefit the city through reducing congestion and pollution, as well as the time taken to complete a journey.[20]

Boston has developed an app that uses the motion sensor in mobile phones to identify when a car hits a pothole. This has allowed the city to identify and solve road maintenance issues much more quickly than through the use of traditional road inspections (case study 6).

---

20 http://www-03.ibm.com/press/us/en/pressrelease/41068.wss

### Case study 6: Boston, USA

The city has developed Street Bump,[21] an app that uses mobile phones to map potholes in the city. This uses the motion sensor built into phones to recognise when a car hits a bump in the road. This data is then transmitted along with a location reference taken from the phone's GPS. Existing speed bumps are already mapped so they don't get mistaken for defects. If a number of people hit a bump in the same spot, the system recognises it as a pothole and the city inspects the site to determine if repairs are needed. The app is free to download and aims to create a real-time map of road conditions to catch problems earlier than traditional inspections.

But as the Santander case study shows (case study 7), the use of sensors to collect data has led to city benefits in areas beyond transportation. The 12,000 sensors placed all across Santander have led to a range of improvements, including a reduction in the city's utility bills, as the sensors allow street light to automatically dim when no-one is around as well as alerting refuse collectors which bins need emptying.

### Case study 7: Santander, Spain[22]

The city has received funding from the EU to become a prototype for smart cities across Europe. Due to its small size, the city can be monitored by 12,000 sensors, providing an opportunity to assess the benefits smart cities can deliver. The sensors collect measures of air pollution, noise levels, identify available parking spaces, inform refuse collectors which bins are full and automatically dim street lights when no one is around. As well as sending this information to a control centre, residents receive real time information via their smartphones on issues such as road closures, parking availability and bus delays. Residents are also able to report issues (such as broken streetlights) directly to city hall. The city has seen a reduction in response times to addressing problems and a 25 per cent fall in electricity bills. The city's rubbish collection costs have also fallen by 20 percent. Due to the savings seen, utility companies are happy to pay for the sensors' upkeep.

**It is not just cities in developed countries that are looking towards big data, cities in countries such as the Ivory Coast and Brazil are also utilising big data.**

IBM analysed the mobile phone data of Orange customers in Abidjan to develop a solution to the city's congestion problem that reduced average journey times by 10 per cent (case study 8).

### Case study 8: Abidjan, Ivory Coast

In 2012, Orange released 2.5 billion call records from five million mobile phone users in Abidjan, the Ivory Coast. Using location data from 500,000 of these phones IBM analysed congestion in the city. The result was a model that predicted that the addition of two bus routes and enhancement of a third would improve the city's transport system, resulting in time savings of 10 per cent for users. Before release, Orange removed any personal information from the data that could be used to identify individuals.[23]

---

21 http://streetbump.org/
22 http://www.npr.org/blogs/parallels/2013/06/04/188370672/Sensors-Transform-Old-Spanish-Port-Into-New-Smart-City and http://www.smartsantander.eu/index.php/testbeds/item/132-santander-summary
23 http://datasmart.ash.harvard.edu/news/article/cell-phones-bus-lines-and-privacy-in-cote-divoire-243

Rio de Janeiro is using data from sensors to connect the whole city, with the information fed back into the city's control centre. This is enabling more co-ordinated and quicker responses to crises, such as building collapses (case study 9).

### Case study 9: Rio de Janeiro, Brazil

In 2010, the city asked IBM to create a city operation centre using data from cameras and sensors located throughout the city as well as connecting all of the city's 30 agencies, from transport to the emergency services. The centre is enabling officials from across the city to collaborate to manage the movement of traffic, while also ensuring that power and water supplies work more efficiently. The city is also now better placed to deal with a crisis, such as a collapsing building, as the system makes it easier to roll out a coordinated response – transport systems can be shut down, emergency services mobilised and gas supplies cut off, while individuals can be informed of alternative routes via Twitter.[24]

These case studies provide an indication of how UK cities could apply big data to improve their own performance.

**To date, UK cities have lagged behind in their use of big data to address city specific issues.**

More recently progress has been made. In 2012 the Technology Strategy Board (TSB) ran a competition asking cities to submit proposals setting out how existing technology could be used to integrate city systems. Following the competition £24 million was awarded by TSB to Glasgow in 2013 to be a Future Cities Demonstrator. The use of big data features heavily in Glasgow's plans. The city will create an open data platform that will provide real time information on transport disruptions, hospital A&E waiting times and energy use throughout the city, potentially identifying ways to lower energy use and reduce fuel poverty (case study 10).

### Case study 10: Glasgow, UK

In January 2013, Glasgow City Council was awarded £24million from the UK Government's Technology Strategy Board (TSB) as part of its City Demonstrator project. The funding will be used to develop a City Management system that will provide an open data platform, improving the functioning of the city in a number of ways. In particular, the council aims to reduce overcrowding on public transport and congestion on roads, thereby lowering pollution, by providing people with real time information on traffic levels on roads, bus and train delays and the location of nearby empty parking spaces. They also hope to reduce A&E waiting times by providing real-time information on waiting lists in hospitals around the city, allowing individuals to identify which hospitals would be able to see them first. Energy levels across the city will be monitored, including the new Combined Heat and Power (CHP) systems which will allow the city to store energy when demand is low and then use it during times when demand is higher. This has the potential to cut people's fuel bills and so help reduce fuel poverty. Through monitoring footfall and retail demand the specific areas of the city attracting visitors will be identified. Combined, all this information could be used to generate a quality of life index for the city.[25]

*"The UK has lagged behind in its development and adoption of big data to address city specific issues"*

24 http://www.bbc.co.uk/news/technology-22538561
25 http://www.glasgow.gov.uk/index.aspx?articleid=9647

TSB also provided Peterborough, London and Bristol with £3million each to develop certain aspects of their proposals. Bristol is using its funding to link data from universities, SMEs, and residents to discover how it could be used to improve city performance. Those that have supplied data will be able to trial any new products and services developed using the data (case study 11).

> **Case study 11: Bristol, UK**
> In April this year, Bristol City Council received funding from the TSB to develop its Citywide Living Lab. This will combine data provided from a range of sources including universities, SMEs, and citizens who have agreed to share their own data. Individuals and organisations will be engaged in using the data to develop new products and services through the staging of hack events and will be able to trial any new products and services.[26]

*"Only those cites who have received funding from a third party are launching projects to utilise big data"*

The case studies also show that, whilst the open data agenda is leaning towards releasing data for which demand is known, some big data projects have taken a more speculative approach. Rather than determining the demand and use for the data before it is released, organisations are releasing the data they collect as part of their day-to-day operations, after removing any personal details. Organisations are also working together to collect new data, interlinking it and then exploring what it shows about the city's performance and future development needs, such as the pollution level data collected in Santander.

Despite the benefits that big data can bring, the use of such data at a city level is still very much in the early stages of development in the UK and it is only those cities that have received funding from a third party such as TSB that are launching projects to utilise big data.

This suggests that activities that emphasise the city level benefits from using big data and the savings it can generate in the long-run, for example through improving existing systems without the need to invest in any major re-developments, are needed.

---

26 Bristol City Council (November 2012) Connect Bristol Feasibility Study.

## Section 3: Obstacles to overcome

While there are clear benefits to be gained from employing an Open Data Policy at both a government and organisational level, there are also concerns that can either prevent the release of data or hinder its release. The case studies illustrate that while some cities are happy to make their data publicly available, many are not.

Making big data publicly available can deliver significant benefits, but there are legitimate concerns associated with making any data publicly available. As outlined in section 1, a lack of demand for data can lower the benefits seen from making data freely available, which can reduce organisations' commitment to an Open Data Policy. In addition, concerns around privacy, cost, continued availability and lack of skills to make sense of vast amounts of data present barriers. These are explored below.

*"Concerns around privacy, cost, continuation of supply and lack of skills to make sense of vast amounts of data present barriers"*

- **Privacy:** Public and private organisations are wary of allowing data to be published if it is possible to identify individuals directly from the data or by linking datasets. However, social media sites such as Facebook suggest that, in specific circumstances, people may be willing to trade privacy around certain issues for access, provided they understand the benefits of giving up their privacy. As such, private organisations may be prepared to publicly share their own data if it means free access to the data of others, or the ability to have free access to any apps developed using the data for a set period of time.

- **Cost of publishing data:** There is a time and financial cost to ensuring data validity before release and presenting the data in a user friendly format. New York has reported that it is concerned it is spending resources organising data for release that no one is using.[27] While this is a valid concern, it is likely that in the UK the benefits of releasing data are underestimated. While the UK is able to measure the costs of publishing data, there is no model for estimating the benefits from re-use of data.[28] Significant amounts of time and money do not need to be spent ensuring data is complete before it is published – publishing incomplete/unverified data may actually help improve the accuracy of data, as individuals and industry experts are better placed to identify and correct local/sector inaccuracies. As the hack events described in case study 4 illustrate, it is not necessarily dependent on the data holders themselves to develop user friendly ways of displaying information – individuals in other organisations can step in and create apps, which further reduces the cost of publication.

- **Uncertainty over continued availability:** In order for organisations to make use of data, some assurance that the data will be updated on a continuing and regular basis is generally required. If user believe the release of data is a one off they are less likely to use it.[29] To overcome this problem the Open Data Institute (ODI) has developed Open Data Certificates.[30] One criterion for being awarded a Certificate is the frequency with which the provider has committed to update the data. While there is no legal requirement for organisations to certify their data, those that do and achieve an award of 'standard' or above provide potential users with the confidence that the data will be published regularly.

---

27 http://techpresident.com/news/23595/nyc-marks-one-year-open-data-law-and-looks-ahead
28 BIS (15 May 2013) Shakespeare review: an independent review of Public Sector Information
29 Centre for Cities interview.
30 https://certificates.theodi.org/

- **Lack of specific skills:** A large amount of skill is required when working with vast amounts of data, particularly when linking various datasets together and presenting this data in a usable format. BIS highlighted that currently there is a skill shortage in the UK and that the lack of data scientists may prevent the full value of open data being realised.[31] However, moves to encourage investment in data linking projects are underway, with the Economic and Social Research Council announcing in April 2013 that it will provide funding of £64million to projects that link up government data, survey data or organisational data in a secure way and/or develop apps that present this data in a user-friendly manner.

Many of these obstacles could be mitigated by increasing the demand for data. In particular, increasing the use of data will increase the benefits of releasing it and lower the average cost of publishing data. As such, organisations that see their data is being regularly used are more likely to continue to publish it. A greater use of data may also lead to more demand for data scientists, more training courses in this specialism and so an improvement to the UK's skill levels.

*"Organisations that see their data is being regularly used are more likely to continue to publish it"*

---

31 BIS (May 2013) Market Assessment of Public Sector Information.

## Section 4: Conclusions

To improve city economies and make cities better places to live and work, more city level data is needed. This will allow identification of the opportunities and weakness individual cities face, which in turn will allow for more effective and efficient policy development to improve the way in which cities function and deliver economic growth.

As the case studies show, open data and big data can be used to deliver a range of benefits, from reducing congestion and pollution to improving service response times and safety. However, currently the UK is not fully exploiting these methods.

The Government holds a range of data that could improve city performance but which is not currently made publicly available. While the Government's move towards an open data agenda is welcome, it needs to do more to ensure the highest possible benefits from pursuing such a policy are realised. This isn't just about the release of more data. It's also about encouraging its use through engaging with organisations and individuals to promote awareness of what data is available. As illustrated in the US, Hack Days that offer funding to turn strong proposals into practical projects and services can be a good way to raise this awareness.

Cities are also in a position to improve their own functioning and to reduce their costs by doing so. Organisations within cities, such as universities and private companies, already collect information that if shared and linked together could give new insights into the challenges faced by cities. Cites can also organize to collect their own data, through, for example, the use of sensors. Despite this, it is only cities that have received funding from a third party that are making use of city level big data in the UK.  Emphasising the savings big data generates in the long-run through improving existing systems without the need to invest in any major re-developments, may encourage uptake.

The UK needs to learn from other countries and expand its investment in and use of government data and big data. This will generate cost savings and allow for the development of targeted and efficient policy, improving the operation of cities and UK economic performance.

*"The government holds a range of data that could improve city performance but is not made publicly available"*

**Policy recommendations**

In order to improve understanding of cities:

- Central and local government should prioritise the release of data they hold on city economies in order to provide cities with the information needed to enhance their economic performance and meet the needs of their citizens.

- Central government and cities should set up a series of 'Hack Days' to encourage new uses of the data that they are releasing and secure funding from sponsors to turn ideas into functional products.

- Cities should set up special interest groups to explore what existing public datasets can be used and combined to provide new insights, savings or proto-types reflective of their city's needs.

- Cities should look to share the big data they already collect by setting up on-line platforms that allow organisations to upload their own data and use the data of others.

- Cities should lead on collecting new data in their own economies and identifying uses of big data that could lead to long-term savings.

*"The UK needs to learn from other countries and expand its investment in and use of government data and big data"*

## Authors

Maire Williams is a Researcher at the Centre for Cities. Contact Maire on: m.williams@centreforcities.org / 020 7803 4317

## Acknowledgements

**Supported by Future Cities Catapult**