City data survey report

for BSI in support of understanding data requirements and standards for smart city initiatives

First published in the UK in 2015

by
BSI Standards Limited
389 Chiswick High Road
London W4 4AL

Contents

1. Overview 4
2. Introduction 7
3. Approaches to smart cities 8
4. Conducting the survey 11
5. Assumptions 12
6. Key findings 14
7. Analysis 16
8. Conclusions 24

Appendix A – References and case studies 26

Appendix B – Survey questions and results 29

1. Overview

The British Standards Institution (BSI) is running a programme of work exploring the data considerations that will help establish a decision-making framework for smart city environments. The full scope of this work is being finalized but includes understanding:

- the culture of sharing data and the commercial implications;
- the types of data required;
- how data is used;
- data quality; and
- the implications of the need for anonymity.

To support this, BSI initiated a project covered by this report to explore data use with the aim of addressing some of the points above. Additionally, it is well understood that data relating to building information modelling (BIM) is inextricably coupled with data relating to smart cities and therefore we ensured this topic was included and we have made reference to this in our report.

For this study, BSI invited Ordnance Survey (OS) to assist in this project to conduct interviews and run a survey with smart cities to address the above points.

Working with a group of 30 UK and international cities, we aimed to identify the most important and relevant datasets through understanding the challenges that, cities had and the projects being initiated to address these challenges. Where possible we aimed to obtain a number of supporting case studies demonstrating their uses and benefits. BSI and OS would like to thank everyone who took part in the survey and for providing a valuable contribution to this report.

The primary focus of this report is on data requirements and whilst there is some necessary discussion on the philosophy of smart cities, it is not the purpose of this document to judge any approaches being adopted, or to provide analysis of any of the specific challenges or projects in great depth.

The findings of this study can be summarized in the pie chart (Figure 1) and Table 1. The pie chart (Figure 1) shows the segmentation of datasets required to support work and initiatives defined by 14 categories:

Required datasets to support city projects

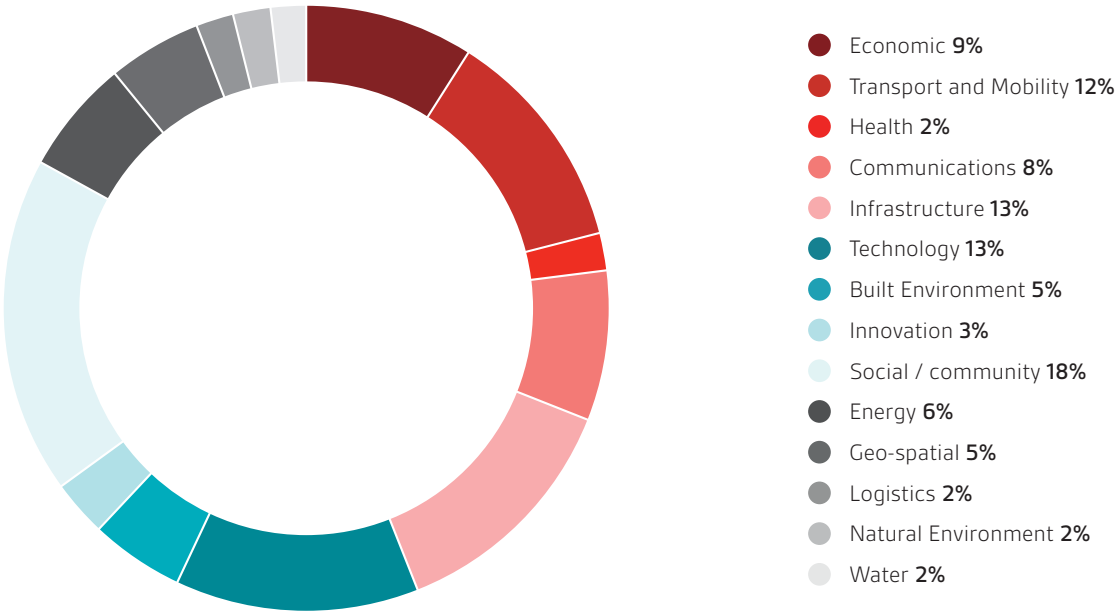

Figure 1 – Required datasets to support city projects

What this chart clearly shows is that the four most prevalent needs for data are those which support:

1. Social and Community.
2. Transport and Mobility.
3. Infrastructure.
4. Technology.

Arguably data relating to BIM shown as 'Built environment' could be classed as Infrastructure but the nature of the responses obtained allowed this to be separated out.

Table 1 summarizes the responses to some of the other key questions raised.

Key survey findings	
Who benefits the most from resolving the challenges within cities?	1. Residents 2. Economy 3. Local government
Proportion of open data being used:	1. Data you own (collected) **47%** 2. Open (free) data **31%** 3. Commercial (paid) data **22%**
Current largest sources of city data:	1. Local government 2. GIS providers 3. Real-time sensors
Barriers to obtaining data:	1. Anonymity – Removal of personal information from datasets 2. Competition – Utilities releasing customer and performance information 3. High costs to obtain, e.g. Mobile phone operators data for footfall and user profile 4. Costs of technology – Creating costly projects using sensors 5. Silos – Obtaining information across government departments within councils
Proportion of the data, cities have or will be using from available sources:	1. 55% Government 2. 16% Sensor technology 3. 16% Crowd-sourced 4. 15% Private sector
Datasets cities would like to use more in the future:	1. Mobile phone operator data 2. Crowd-sourced data.
Which of the following challenges are also relevant to your city?	1. Mobility / Transportation 2. Traffic congestion 3. Business generation and energy conservation

Table 1 – Summary of responses to key questions

It is recognized that cities have both diverse and common challenges and applying standards vertically may be ambitious. A more horizontal approach to establish a framework can be established and the four categories identified from Figure 1 will provide a good initial focus.

Expectations are that data should be open but it is recognized that this does not mean it is necessarily free. Work needs to take place by all parties to understand the common benefits, sharing data will bring to each. This will require a better understanding of value and consideration for a commercial framework to exist within and between cities.

As a major benefactor from smart city development, the citizen will have an increasingly important role to play through providing crowd-sourced data allowing smarter and quicker decisions to be made.

Data relating to the economic and financial factors will be required to assist in decision-making.

1.1.1 Building information modelling (BIM) relationships

During our conversations with some of the city representatives it became increasingly evident that the relationship between smart city data and that required for BIM is brought together by the need to share data and information. The information architects' and engineers' needs are similar to those required for sustaining a smart city although historically this information has been exploited solely for specific project build, planning purposes and development. Historically new building projects have considered the relationship with the surrounding environment but only to a limited extent.

In developing and maintaining a smart city, the surrounding environment and the long-term impact requires data to persist and be readily shared to allow assessment of the impact and interaction between buildings and the city as its 'landscape' and needs change over time.

Ensuring this data is available requires all parties to have a shared vision resulting in shared benefits.

Cities across the world publish several hundreds of datasets, and we can expect this number to rise rapidly in the future.

BIM and smart city development are inextricably linked indicating a shared need for common datasets. Given the strong and necessary relationship between smart city and BIM, BSI already has a number of projects underway to explore this.

2. Introduction

The project was conducted in two phases and this report covers the findings from both Phase 1 and Phase 2.

For the first phase, OS interviewed 11 city representatives working with smart city data to get a preliminary view of what they think the core datasets for a city should be, the potential uses of this data and the perceived benefits of using it. The contacts were made with:

1. Barcelona (Spain)
2. London (England)
3. Bristol (England)
4. Leeds (England)
5. Dubai (United Arab Emirates)
6. Peterborough (England)
7. Glasgow (Scotland)
8. Sedgemoor (England)
9. Singapore (Singapore)
10. Beijing (MOHURD)[1] – (China)
11. Greenwich (England)

Phase 2 made use of an online survey approach to enable us to reach out globally to a wider group of cities. This survey design and approach built upon the findings from the Phase 1 work. Whilst contact was attempted with more than 125 cities globally, only 30 cities responded of which 20 chose to participate? These are:

1. Aberdeen (Scotland)
2. Santander (Spain)
3. Ghent (Belgium)
4. Vienna (Austria)
5. Preston (England)
6. Tampere (Finland)
7. Cardiff (Wales)
8. Munich (Germany)
9. Israeli Ministry of Energy
10. Copenhagen (Denmark)
11. Perth (Scotland)
12. Rijeka (Croatia)
13. Valencia (Spain)
14. Coruña (Spain)
15. Malaga (Spain)
16. Bordeaux (France)
17. Birmingham (England)
18. Stockholm (Sweden)
19. France (Cerema.org.)[2]
20. Perth (Australia)

In Phase 1 we developed a briefing document to position the work. It defined seven questions to be covered during the interview. We spent about 10 days locating and contacting each city representative and then held a telephone interview with each one to address the questions posed. Each interview was captured on a template ensuring consistency of approach.

The interviews proved challenging in a number of ways, with obstacles including: language differences, time constraints and zones, maintaining focus, concerns over the level of detail required, the lack of detail offered, the embryonic state of some cities' engagement, the diversity of these engagements and a lack of case studies.

In Phase 2 we worked with a leading online survey company 'Questback'3 using their 'easyreach' tool to develop a detailed questionnaire comprising 22 questions. Identifying the right contacts and seeking agreement to participate was a challenging piece of work. In addition four cities took longer to respond and missed the agreed deadline and as a consequence we have not been able to include them in the survey results. Not all participants completed all the questions which meant some assumptions had to be made and the data 'weighted' to reflect this. The detailed results of the survey itself are found in Annex B.4.

[1] Ministry of Housing and urban-rural development http://www.mohurd.gov.cn/

[2] Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement http://www.cerema.fr/

[3] Global online survey provider http://www.questback.com/

3. Approaches to smart cities

In contacting the cities the results from both Phase 1 and Phase 2 has revealed that each city is at a different stage of its development. The spectrum is wide-ranging from those who are considering the next 50 years plus and developing a vision, to those who are developing strategies and addressing their funding issues, to those who are beginning to solve real city problems now in smart ways and those who are more focused on innovating and exploring capabilities of new technologies. What this means is that an equally wide spectrum of data is needed to support decision-making. For example, those with visionary aspirations need to use varieties of data to which they can apply modelling techniques to predict outcomes whilst those solving problems today need current and real-time data to make daily decisions for the smooth operation of the city.

3.1 Two models

Whilst obtaining results from just 30 cities does not provide a statistically accurate set of global data for full analysis, the work has revealed that in general, there are two ways in which cities tend to determine what data they need and how to get value from it.

Firstly, there are those who seek out problems in their city environment, whether it is addressing social, health or transportation issues and then look at the data they need to build solutions to address these. This is a logical approach which delivers benefits, but could be viewed as being a little cautious, too focused and considered. Cities just starting out to develop strategies for their smart city are also met with political and financial challenges and these add additional pressure throttling the ability to make rapid headway and leading to a more cautious approach.

Secondly, there are those who recognize you can acquire data from many sources, both static (such as published statistics) and dynamic (such as traffic flow or daily footfall) and they make every effort to exploit this by consuming vast sets of data and by deploying sensor technology to acquire real-time data, even if it is not fully understood how this will be used. This approach is more ambitious, visionary, and opportunistic.

These cities would also appear to have a greater degree of funding available although this was not fully explored during this study. However we did ask about budget for data but responses were guarded and those who did respond suggested around $100,000 would be put aside for this.

The two approaches can be illustrated as shown in Figure 2.

In reality those cities making good headway are adopting both of these approaches in a balanced way recognizing the need to be both visionary but with a focus on the benefits and costs too.

During the survey there was also a set of responses that focused much more on the ability to monitor and control the city and therefore a need for data sources that can provide timely feedback. In particular some cities in Spain and Singapore and Beijing have advanced control centres. These cities are already reasonably mature in having addressed some of the key problems in their city and now need data to monitor and improve performance to continue to deliver and sustain the vision for their city.

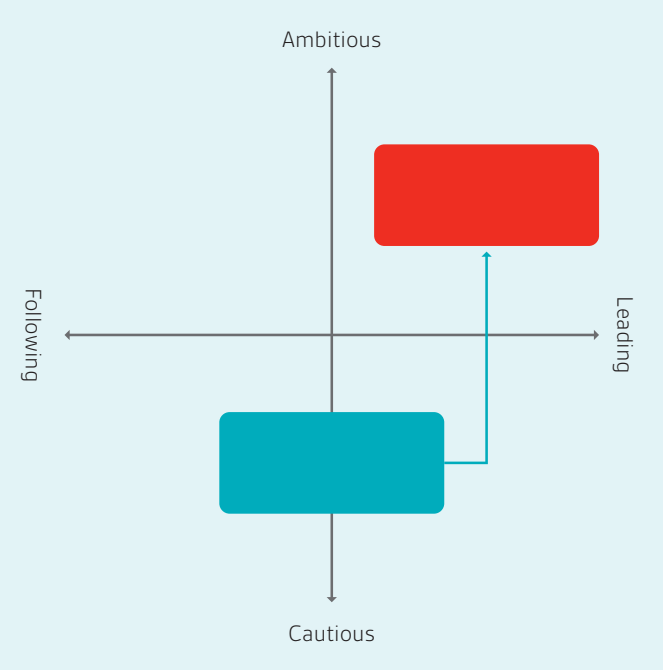

Figure 2 – Smart city approaches

3.2 BIM Considerations

Cities have a built environment embracing many types of buildings, structures, transport routes and natural geographical objects. How these interact with each other has an impact on the city operation in terms of how individuals and communities exploit their purpose, make choices and decisions as to where they live, work and socialize as well as determining how they navigate the city landscape. In a smart city, future planning and adapting to rapid change will inevitably require new developments to be built or existing developments to be modified, upgraded or demolished depending on the needs of the city and its communities. During planning of such change, the interaction and impact needs to be understood before costly (in the broadest sense of the word) decisions are made. Traditionally when new developments have taken place, information is pooled from a variety of sources to manage the build through its lifecycle. However it is not uncommon for this information to be 'closed' in its nature and is used for the sole purpose of completing the work and is not widely made available to other interested parties.

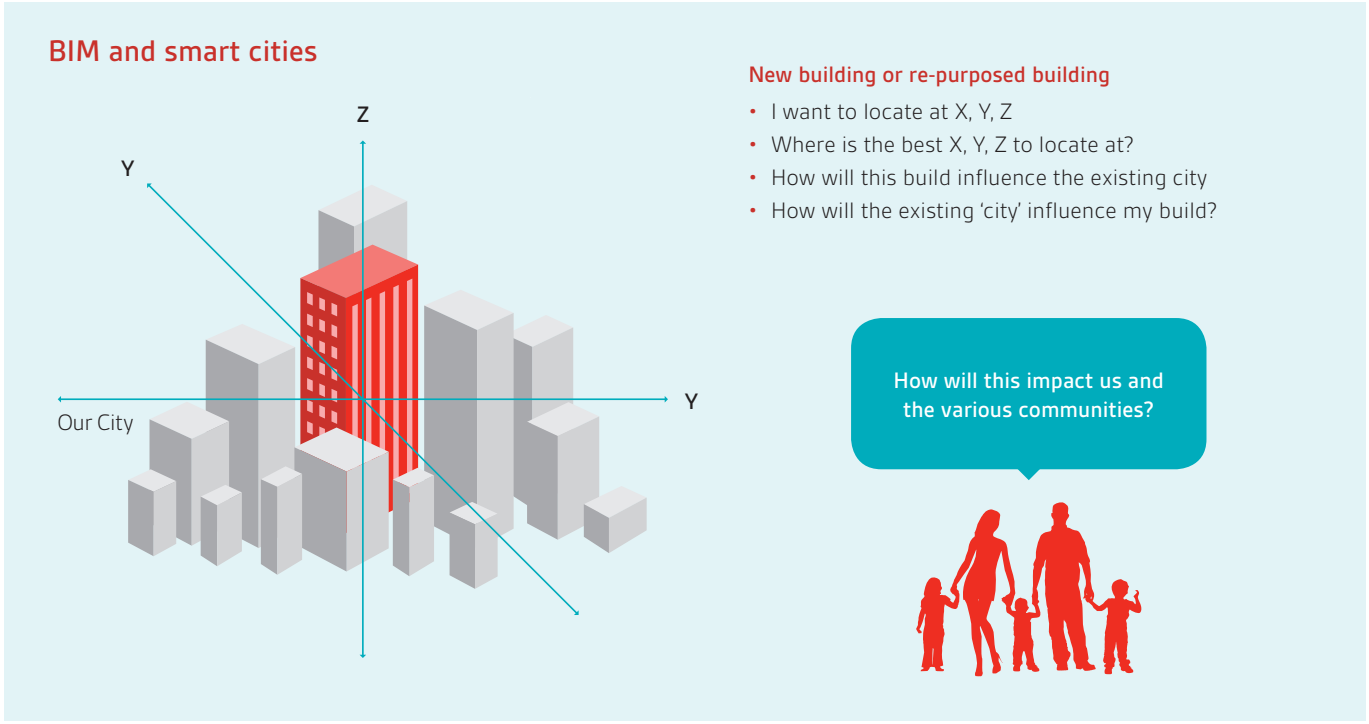

Figure 3 – BIM and smart cities

In the smart city environment we have learned that expectations are that data needs to be open and freely available for whoever needs it to enable the city to continuously develop and improve. This means that data or information relating to new developments also needs to be open to allow the completion of the full picture. Civil Engineers have indicated that they would not expect to be constrained by lack of specific datasets so it seems reasonable that equally the communities within a city should also not be constrained by developers and construction companies withholding information before, during and after builds.

BIM recognizes that information needs to persist through the whole lifecycle of development and needs to be made available or shared to a wide range of communities including the asset owner, financiers, developers, architects, residents and local council to name some. By providing this information or data in a BIM environment it becomes possible to model the impact and interaction on for example, transport demands, healthcare, social requirements, green space, emissions / air quality, retail activity and, informs decisions on where to site new or modify existing developments with minimal adverse impact on the city at large and the communities within.

Smart cities and BIM are clearly connected as one has a profound impact on the other and the decisions that hang of them. Ultimately for both to thrive, they cannot be dealt with in isolation and it is easy to appreciate that shared information and shared data will provide the platform needed to enable both to coexist and grow. Ageing population is an increasing global problem. People are living longer and need to be housed and receive healthcare and have access to suitable mobility choices. Designing cities to cater for this requires close cooperation.

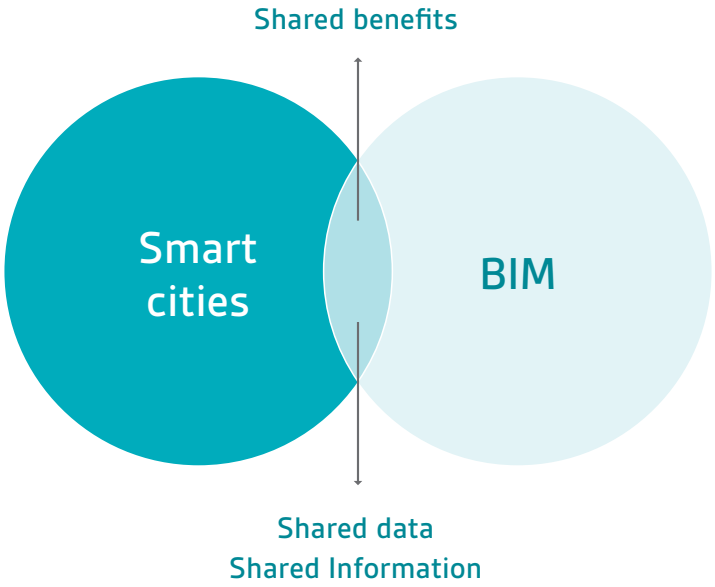

Figure 4

BIM embraces 3D visualization which greatly assists in understanding the impact of developments allowing the key decisions to be made with a higher degree of certainty. This also provides a greater degree of accessibility to the affected communities to observe the potential impacts and feed in information during the planning stage providing a much greater chance for a successful outcome.

BIM helps bring together the Geospatial and the Construction design and engineering world to create this powerful modelling environment that everyone can be involved in, challenge, and contribute to. Modelling cities using similar resources is a natural progression. The use of 3D Geospatial data will inevitably increase allowing greater and more realistic 'what if' modelling.

4. Conducting the survey

4.1 Phase 1

Over a period of 10 days, we aimed to make contact with a number of cities that were identified as having a significant interest in the development of smart cities and were likely to be able to provide information on the use of datasets in a smart city environment. Eleven cities took part along with one civil engineering company who we used to seek their views from a BIM perspective, giving a total of eleven contributors to the Phase 1 study.

Each contact agreed to be interviewed by OS initially for thirty minutes but a small number were prepared to talk about their experiences for up to one hour. On reflection we have concluded that thirty minutes is probably the bare minimum time required to search out meaningful information. During the call we aimed to cover seven questions (see Annex B.1).
The bulk of the discussion focused mostly on the first three questions.

i. What datasets do you consider essential for developing a smart city environment?
ii. Where you expect to obtain these datasets – open source, internal or other?
iii. Which datasets are you already using and for what purposes?

In general, question (i) was deemed too broad to answer and it was therefore necessary to 'seed' the question with additional information such as 'have you thought about datasets relating to…?' The problem with this approach is that it is then very easy for the interviewee to simply say 'yes' rather than thinking about those sets of data which are really important and are genuinely adding value to their city challenges. We therefore tried to avoid prompting where possible unless the conversation was stagnating or not moving in the desired direction.

4.2 Phase 2

Having understood the challenges from Phase 1, a modified questionnaire was introduced suitable for completion online. The biggest challenge was to secure representatives from global cities who were willing to participate and who were informed about the data needs. We used several resources to achieve this including *https://eu-smartcities.eu/*, internet searches, BSI and OS contacts, the 'smart-to-futures cities 2015' event and the cities contacted in Phase 1.

In all we attempted to contact more than 125 cities but those who chose to participate were just 20. The survey comprised 22 questions (see Appendix B.3) and was designed to focus the participant on thinking about the problems their city has to then drive out information regarding the datasets that would be needed to support resolving these problems, and the sources of this data.

The survey was open for three weeks[4] to enable each city adequate time to complete the 20 minute survey. The survey was hosted by an external organization 'Questback' providing an impartial platform to receive the results.

Not all participants completed all 22 questions which has made the analysis a little more challenging than anticipated but appropriate weightings during analysis to the responses have been applied to take this into consideration.

[4] A future survey would benefit from being undertaken over a longer period as it was very difficult to secure time of our valuable contacts that were in high demand to respond in a timely manner. We also found a small number of cities did not respond until over a month later but would have participated given more time.

5. Assumptions

During Phase 1 all the information gathered was collected verbally and then documented as free text. The text was analysed to identify and distil specific themes and categories of data and these were referenced against what are fairly well established building blocks and frameworks for smart cities. This initially identified 10 categories we could reference the data against.

In Phase 2 we asked the participants to identify the support datasets explicitly which made aligning these two categories within the framework a little easier. Because the responses were 'free text', the quality of the responses was wide ranging and it was necessary to make educated assessments of the categories that each dataset might fall into. Due to the wider set of questioning in Phase 2 and by focusing on specific projects, it was clear that a further four categories (in **bold**) should be added to those identified in Phase 1. The final 14 categories are shown below:

1. Infrastructure – KPI measures, Commercial, industrial activity measures.
2. **Built environment – Assets, buildings, planning data, buried utilities**.
3. Energy – Electricity, gas, renewables.
4. Water – Supply, demand, natural flow and flooding.
5. **Economic – Financial costs, financial benefits, financial asset flow, growth**.
6. **Geo-spatial – Location, routing, terrain**.
7. **Innovation – Ideas generation, creative thinking, new concepts**.
8. Logistics and services – City services such as waste management, street maintenance.
9. Transport / Mobility (public and private) – Road, rail, pedestrian, tram, bus, cycle.
10. Natural environment – Weather, climate, air quality, green spaces, trees.
11. Communications – networks (e.g. fibre, wifi, cellular), access points, Internet of things (IoT)
12. Social – Communities, deprivation, jobs, skills.
13. Health – Patient statistics, disease types, disabilities.
14. Technology – Sensors esp. real-time.

This framework could be further simplified to:

A. data that defines the community and environment;
B. data that impacts on health and social wellbeing;
C. data that helps with city efficiencies, services and effective use of assets; and
D. data that has to be published to meet political and statutory requirements.

For this survey, the above four categories proved useful as 'prompts' during the Phase 1 interview but the wider fourteen categories proved more useful and relevant for this study. The resultant categories and typical datasets that fall within these categories are shown in Table 1 to the left. In some cases for Phase 2 we have placed datasets into more than one category depending on the nature of its purpose in resolving a the city problem:

Table 1 – Dataset categories

Infrastructure

- Asset location
- Maintenance schedules for assets
- Industrial data e.g. emissions
- Retail data e.g. trading figures, footfall
- Quality indicators

Energy

- Supply
- Demand
- Emissions
- Maintenance
- Profiling e.g. street lighting
- Home insulation statistics

Water

- Supply
- Demand
- Maintenance
- Flood risk

Logistics

- Garbage management and collection
- City mobility
- Transport for services

Transport / Mobility

- Traffic flow
- Vehicle recognition
- Cycle route planning
- Feature identification (drop kerb)

Economic

- Economic participation by residents to make informed decisions based on city financial data[5]
- Gross value added (GVA)
- Pay rates by job
- Spend on services
- Cost of resources
- Economic data as a result of actions (benefits)
- Sector growth

Built environment

- Building types and parameters e.g. business rates
- Building improvements and re-purposing
- Energy ratings
- Planning information

Geospatial

- Geographic features
- Cadastre
- General map data
- Travel routes
- Geographic location
- Terrain

Innovation

- Innovation related data
- Smart city thinking
- Innovation resources

Natural environment

- Climate and weather
- Air quality / pollution
- Green space
- Service improvements
- Biodiversity
- Tree locations (including hazards)
- Land use

Communication

- Networked cities – Fibre optic with access points
- Mobile / Cell phone
- Crowd-sourced e.g. incidents, trends, new data gathering

Social/community

- Crowd-sourced
- Crime statistics
- Fuel poverty
- Domestic living standards
- Social housing
- People movement
- Population
- Communities identification
- Demographics
- Public safety
- Ageing of citizens
- Child care
- Political landscape
- Funding decisions

Health

- Hospital admissions
- Patient diseases / conditions
- Pest control
- Fly tipping

Technology

- Static sensors
- Traffic monitoring
- Weather stations (including air quality)
- Mobile sensors, e.g. public transport and crowd-sourced

[5] Citizens having access to financial information (e.g. dashboards) regarding their city and understanding the choices available allowing them to make decisions based on the economic impact of these choices. For example, how should they travel, how should they recycle, how should they respond to proposals for news developments?

6. Key findings

Specific requirements for smart cities vary depending on demographics, industry, geography and the economy. However, several city challenges are similar and because of this, trends in the responses have identified a core set of problems. Given this, it is not unreasonable to assume that similar datasets will be required globally, albeit to different degrees. The types of datasets used within cities can vary considerably and can run into several hundreds. When examining the challenges that cities have this is where we see some correlation and we are able to group datasets to fit into pre-defined categories. Over both phases of the work we identified 14 distinct categories which would help describe the datasets (see Figure 1) at a meaningful level.

We observed that although in Phase 1 the participants in the telephone survey made reference to the natural environment, this topic was not reflected strongly by the participants in Phase 2.

In Phase 1 we identified 10 categories that adequately described the groups of datasets but as Phase 2 progressed we expanded this to 14. With more cities taking part it became easier to distinguish between Infrastructure elements and BIM elements, draw out the geo-spatial elements and observe a desire for sets of data that would help promote innovation. In Phase 1 there was some evidence of economic considerations, efficiency of services, cost considerations, but in Phase 2 this was more transparent and therefore justified a separate category.

From the information gathered in both phases of the study, there was strong evidence and an expectation that all data required to deliver a smart city capability – both that which is obtained by the city and that which is published by the city – should be 'open data'. It was recognized that whilst this is relatively easy to achieve in the public sector due to statutory obligations to make data publicly available and in the private sector there are a number of barriers that exist and can be summarized as follows:

i. 'What's in it for me, i.e. how much will you pay me?'
ii. 'The data holds personal information so I can't give it to you!'

Although not cited through the survey, OS's experience also identifies that data relating to utilities can also be guarded due to the potentially high risks of exposing data relating to positions, and types of pipelines, ducts and cables.

Sometimes the barriers are not always real and can be overcome. By way of example, one UK city we interviewed in Phase 1 approached a water authority which clearly thought it could not release any data as it held personal information even though they could see the shared benefit. However the council proposed that it could publish its data excluding personal information. The water authority not only agreed but realized by excluding the personal elements there was quite a lot of additional data that could be made available for publication.

In Phase 1 there was some evidence of the need to have data to specifically help manage costs and efficiencies and Phase 2 reinforced this need. This is important as the need to understand costs and benefits associated with projects in the context of a diverse and ever changing city landscape will be put under ever increasing pressure. Availability of data to assess the overall financial and economic position will therefore become significant.

Innovation has become a major factor in many cities future development and in Phase 2 we saw evidence of the need to provide data that supports innovation and creativity. Many smart cities now have developed innovation centres or hubs to foster the creative environment.

Crowd-sourced data was identified as a major requirement in the future. This provides real-time data that enables decisions and action to be made in a timely manner and, can provide instant feedback, which is cost effective, valuable and further engages the citizen.

Data sources are varied with the majority coming from public sector government based web sites and sensors.

BSI is acutely aware of the barriers to sharing data that not only exists between public and private sector but also within the public sector itself and is commonly referred to as the 'Silo mentality'. Both UK and some of the Eastern European countries taking part in this survey and, mention of this at the Smart to Future City 2015 conference re-enforced this awareness.

There is consensus that not only should data be 'open' but all data should be easy to ingest, in general this seems not to be a problem as whilst data can be provided in many formats and through many media channels, it is felt that today's technologies avoids this being a specific barrier to use.

Another UK city was very interested in the value of information from mobile phone operator's cellular location data. However, the mobile operators appear only to want to provide such information at a premium price, and this is seen as prohibitive to usage.

These private sector views may well change if the benefits of sharing data can be demonstrated and appropriate education and awareness of these benefits shared amongst the commercial data providers.

Another major source of data is expected to come from crowd-sourced real-time hand held devices. In most cities there are often more devices than people and access to devices has become increasingly easier over recent years. There is great potential to harvest data from new mobile App's. and the real-time aspect allows cities to make faster decisions to improve the city environment when and where it needs to be improved. This will generate large amounts of data and will be valuable to both private and public sector in shaping the city landscape. This is likely to be one area where the data being acquired may be city specific but if standards are applied to the data schemas then this will aid interoperability within cites and ultimately over time, once embedded, interoperability between cities globally.

Sensor technology is being widely exploited, providing real-time data also allows for faster decisions both with and without human intervention e.g. for modifying traffic light sequences based on congestion, or flood controls to close barriers or divert water courses. As congestion increases, parking sensors are becoming more prevalent and by way of example in Singapore the 'onemap' portal *http://www.onemap.sg/index.html* provides parking availability information as well as a rich suite of information describing the city 'landscape' and activity.

Figure 5 – Singapore's 'one map portal'

The need for standardization and interoperability is a topic that is well understood in the smart cities arena and as smart cities develop, the boundaries both within and ultimately between cities will operationally become blurred allowing us to move from smart city to smart world! Standards are therefore a vital component of the future success of the smart city and the identification of the key data themes in this report aims to provide the necessary focus.

7. Analysis

In Phase 1 we identified 10 categories and Figure 6 shows the breakdown of the datasets groups by the categories in Tables 1 (Section 5), weighted by the occurrence of topics identified during discussion. At this stage we only identified 10 categories due to the small sample of just eight cities which did not provide a statistically significant sample. In Figure 7, we see the results of the larger sample from the 18 cities in Phase 2.

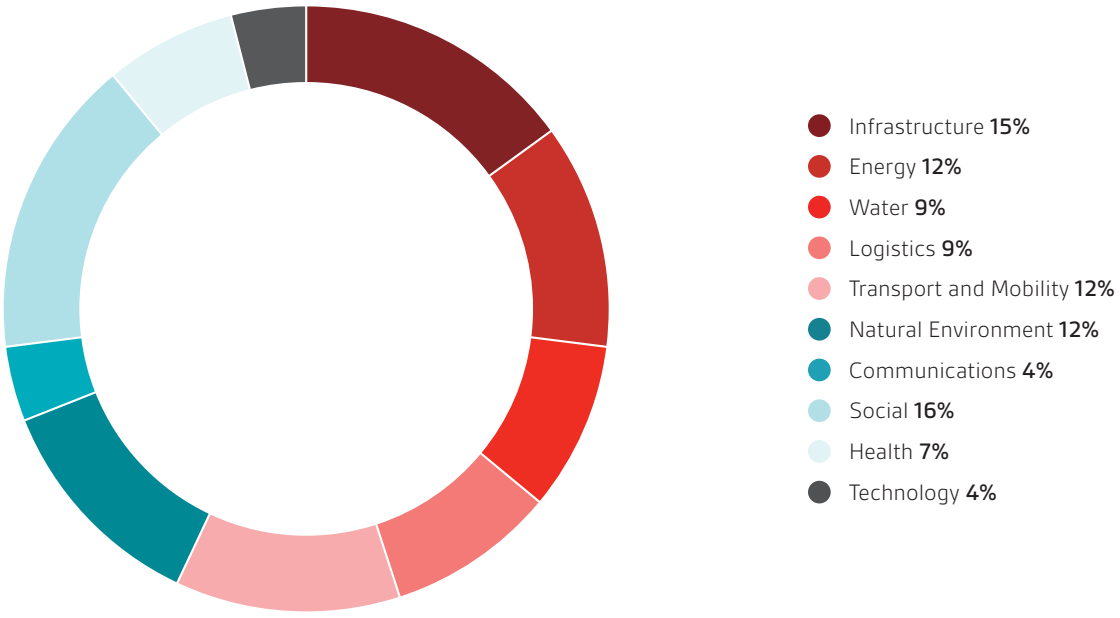

Figure 6 – Phase 1 initial findings

From the results of Phase 1 (Figure 6), two categories are dominant; Social and Infrastructure, suggesting that across the sample of countries interviewed, it is these areas that are the strongest drivers for smart cities.

This conclusion is consistent with the verbal content of responses obtained during interviews.

Required datasets to support city projects

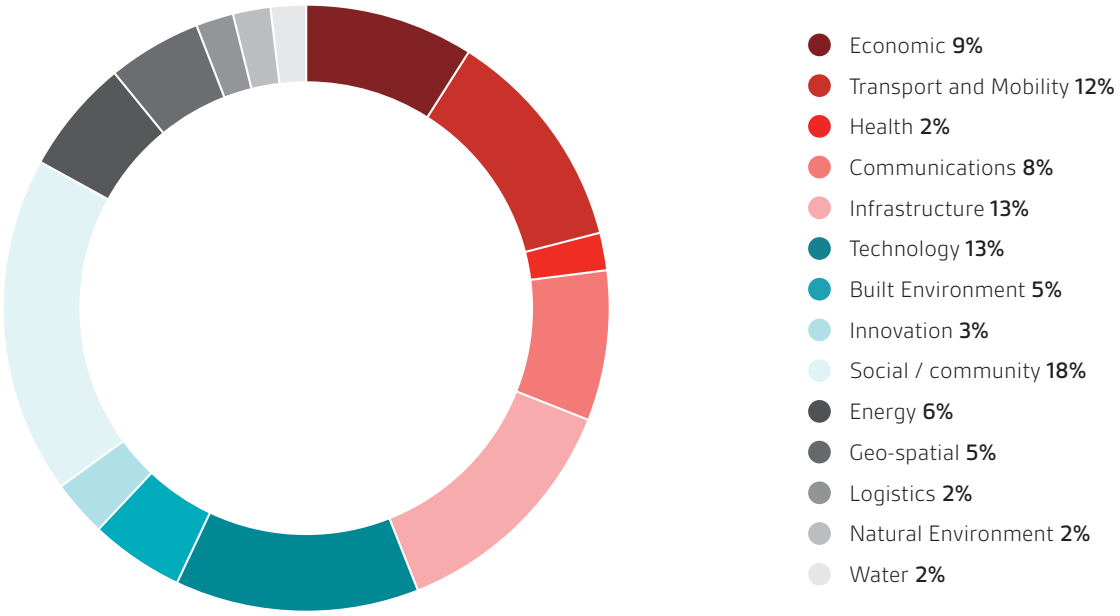

Figure 7 – Phase 2 findings

A direct comparison between the datasets identified from Phase 1 and Phase 2 is shown in Figure 8

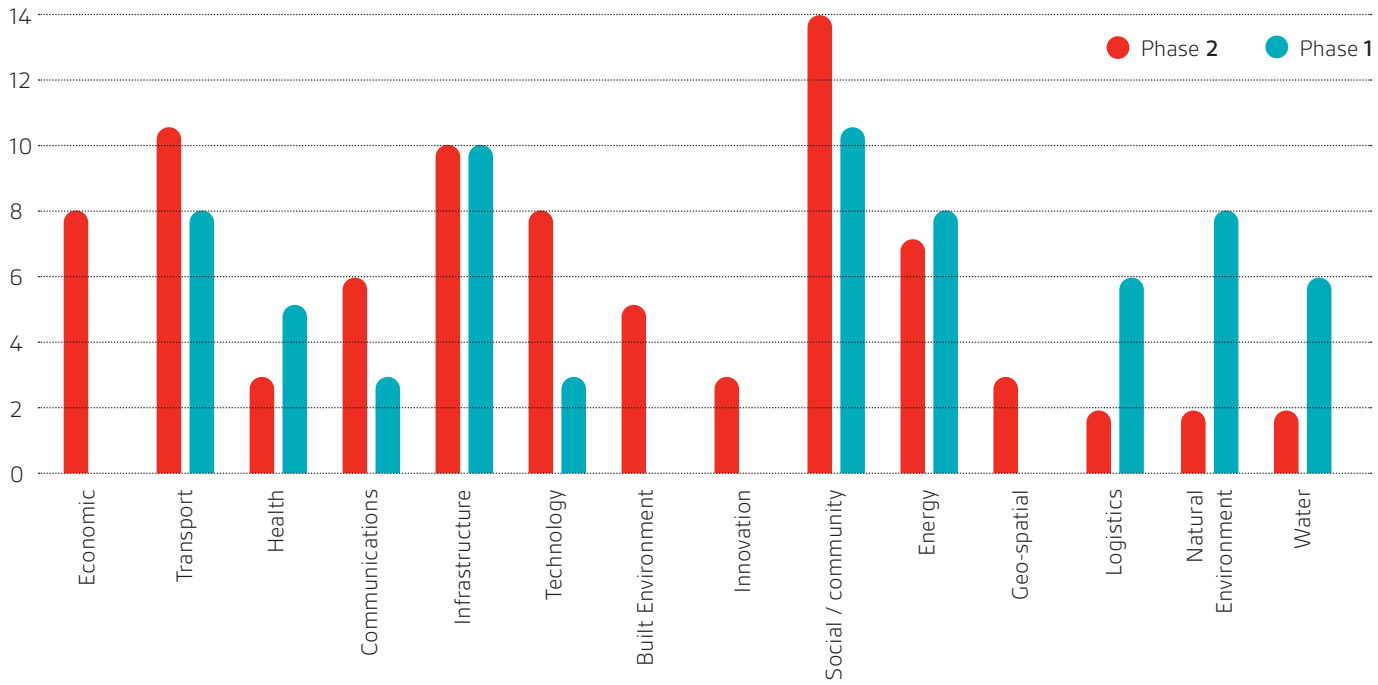

Figure 8 – Phase 1 and Phase 2 comparison

Figure 8 compares the results from Phase 1 and Phase 2 (save the four additional datasets). There is a good correlation between social/community, transport, energy and infrastructure but there are mismatches regarding the natural environment, water and logistics. In Phase 1 the survey was mostly conducted with UK cities and as the Phase 1 survey was done verbally and we were keen to make sure that every category was covered, it may be that during interview, the conversation led the interviewee a little too much by presenting the opportunity to speak about some of the wider city challenges whereas in the online survey we focused participants on the top three challenges facing their cities.

7.1 Published data

Many countries central governments and local authorities are required to publish many of the datasets they hold for their cities as open data. The approach and level of detail will vary from city to city.

Whilst technology is able to manage different formats of data, it is the content of the data, its quality, accuracy, integrity, availability and interoperability that will ultimately limit its usefulness.

In Phase 1 we explored the types of datasets that cities publish, but this was not a focus area for phase 2.

Examples of datasets and their volumes are shown below:

- The Leeds data mill *http://www.leedsdatamill.org/* resource offers an excellent example of more than 150 datasets for download in a variety of standard formats suitably categorized to improve searching.
- Glasgow *https://data.glasgow.gov.uk/dataset* has published more than 370 datasets again suitably categorized and available for download in multiple standard formats.
- Florence *http://www.opendata.comune.fi.it/* has published 850 datasets categorized and available for download in multiple standard formats.
- Helsinki *http://www.hri.fi/en/dataset?q=&sort=metadata_created+desc 1181* datasets categorized and available for download in multiple standard formats.

There is juxtaposition in opening up datasets. Some cities simply make lots of data available to improve transparency but do not explicitly link these to any specific service or community benefit. Some cities also take the view that if they publish as much data as they can then developers and SMEs will make use of this to build useful apps and solutions in support of their cities rather than having to invest in developing these directly themselves.

7.2 Data acquisition

Data can be provided from a number of sources. Typically these datasets are provided through a relatively small number of channels specifically:

- Government datasets published through web portals as open data.
- Government datasets that are not published but reside within government departments.
- Commercial datasets that need to be procured.
- Crowd-source data through mobile device applications or on-line web portals.
- Real-time data sensors such as traffic flow, parking bay monitoring or climate/weather conditions.

Figure 9 shows the relative proportion of data from each source.

Proportion of the data used from each source

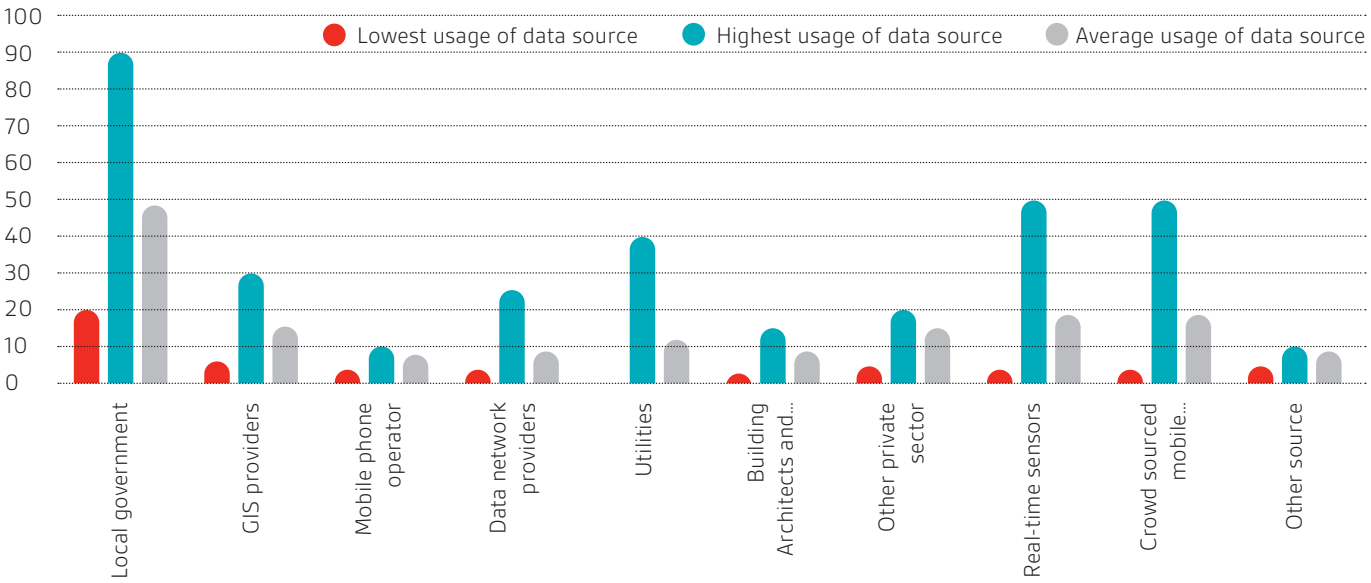

Figure 9 – Proportion of the data used from each source

Use of local government-sourced data is most prevalent followed closely by that provided by real-time sensors and crowd-sourced data.

When understanding the difficulty in obtaining information five areas can be considered:

1. Anonymity – Removal of personal information from datasets.
2. Competition – Utilities releasing customer and performance information.
3. High costs to obtain, e.g. mobile phone operators data for footfall and user profile.
4. Costs of technology – Creating costly projects using sensors.
5. Silos – Obtaining information across government departments within councils.

The last point did not come out directly from the survey but it was prevalent in Phase 1 and was evident through talking to city representatives outside of the survey itself. See 7.2.1.

The delivery mechanisms for data can also vary. Whilst more and more data is made available for direct download from websites in formats such as CSV, some data is being increasingly provided through Application Programming Interfaces (APIs) allowing on-demand ingestion of data whilst other datasets may only be available from media (DVD/CD). The latter is particularly useful for large datasets or where bandwidth is limited but this is not efficient and tends to be updated less frequently as the data has to go through a 'production process' which can be costly and infrequent. For example, some UK map products are supplied on hard media.

For a smart city to adapt to a rapidly changing world, real-time data will be in increasing demand allowing rapid decisions to be made. In particular this applies when concerning transportation and general movement of people reacting to daily changing city events.

7.2.1 Data silos

Whilst government departments often hold the key to providing valuable datasets, unless these are published as open data, it is often difficult for those needing the data to acquire it.

Examples from local councils in the UK suggest that even close colleagues are reluctant to share data. This report does not examine this issue in detail but surmises that often this lack of cooperation is caused through adoption of long-established behaviours and not having a common goal or objective to work towards. Local councils will need to work hard and provide strong leadership to break these silos down if cities are to become smart and capitalize on the wealth of data available.

There are of course datasets that are naturally sensitive[6], in particular any which hold personal information and in those cases, it is common to cite 'data protection' as a reason not to share data and remain siloed. The reality is though, that even these datasets can be made anonymous and still provide high value. For example, health statistics do not need to name individuals but the fact that 1000 people in a certain area have similar heart problems is in its own right, valuable when planning new medical centres, delivering new local pharmaceutical services, understanding the environment, air quality and providing transport and transport links.

The problem of silos is not limited to government. Silos exist between all organizations but often, in the context of smart cities, the silos are created for commercial reasons, i.e. 'how much will you pay me?' and, citing the inclusion of personal information as a reason not to share data.

The enabler to foster data sharing is sharing itself with a clear view of how everyone can benefit. Just as council leaders need to work hard to break down silos, business too has to look at the big picture and understand the wider benefits to all parties.

As volumes of both static and dynamic data grow, the management of this data both proactively and reactively in order to present the impacts to government decision-makers, industry users and the public becomes more challenging.

Some cities are developing city data centres where a wide range of datasets can be presented, analysed and then vital decisions made either by systems or with the addition of human influence to make those decisions. Proliferation of silos will only inhibit this.

Vision and strong leadership are the two vital components that need to be developed to break down the silos and ultimately improve sharing and interoperability.

7.3 Data format

Whilst the drive for the adoption of CityGML is gaining momentum, data used to make cities 'smart' continues to present itself in many other formats, e.g. plain text, spreadsheet formats, shape files, live data feeds (APIs) etc.

In general, discussion from Phase 1 suggested that the variety of data formats was currently less of a problem as most systems are able to convert most files from one format to another with little difficulty. There will always be exceptions but no one identified this as a real challenge to utilizing the data today.

BSI understand that interoperability will become increasingly important and that a data sharing decision-making framework needs to be established to consider data formats and conformity.

Looking to the future data requirements, by examining the responses to Q.17 and 21 (B.4.2), the relative increase in usage illustrated in the Table 2 and Figure 10 reveals a greater use of **mobile phone operator data**, **crowd-sourced data** and data from **Building architects and engineers** is anticipated.

[6] One city reported the need for files which hold extreme private and sensitive data at the lowest level of aggregation. For example, population work needs including individual records containing address, nationality history, parent's nationality history, etc.

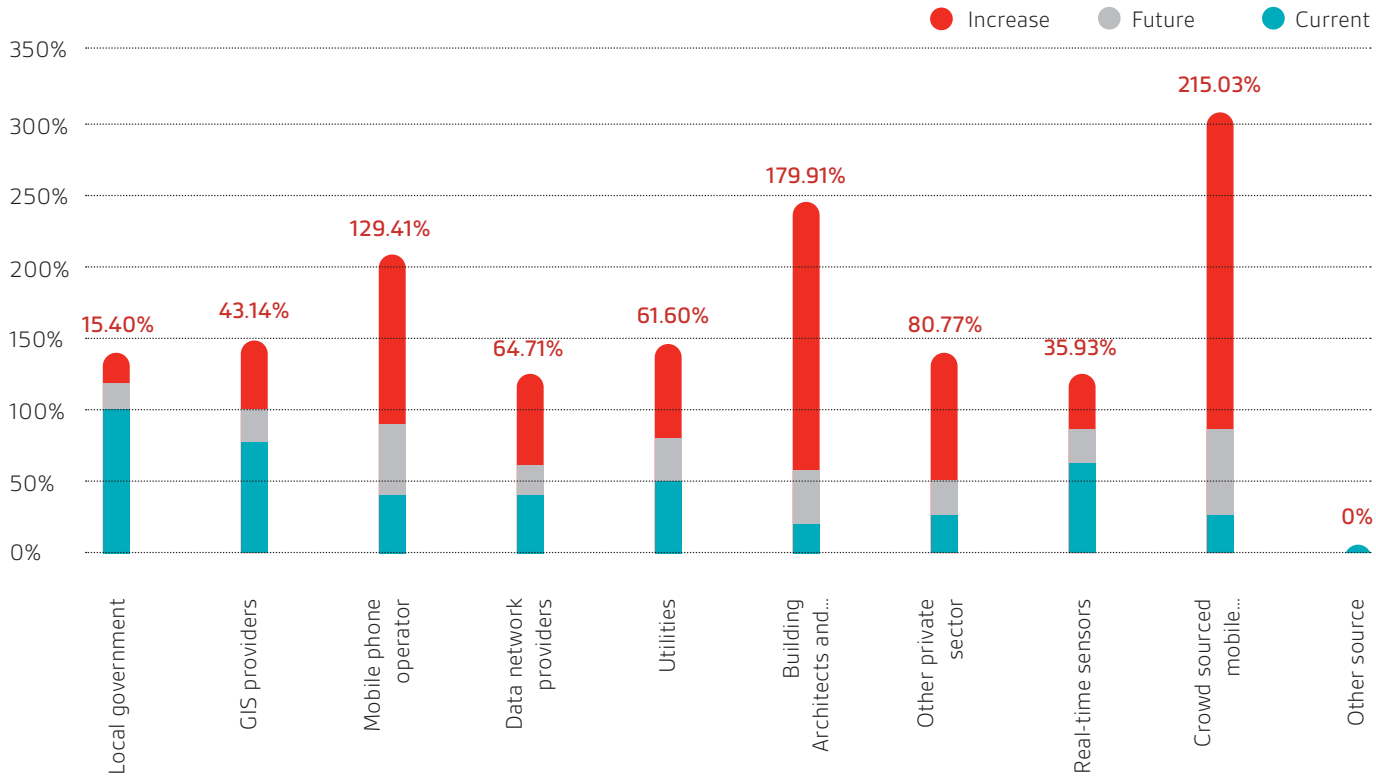

Figure 10 – Data source usage – Current and forecast

Data source	Current	Future	Increase
Local government	100%	15.40%	15.40%
GIS providers	71.40%	30.80%	43.14%
Mobile phone operator	35.70%	46.20%	**129.41%**
Data network providers	35.70%	23.10%	64.71%
Utilities	50%	30.80%	61.60%
Building Architects and Engineers	21.40%	38.50%	**179.91%**
Other private sector	28.60%	23.10%	80.77%
Real-time sensors	64.30%	23.10%	35.93%
Crowd-sourced mobile	28.60%	61.50%	**215.03%**
Other (please specify)	7.10%	0%	0.00%

In summary, all data needs to be able to be communicated effectively and shared with all parts of the city that need to make decisions using common definitions and standards so it can be readily re-used to provide solutions to any specific problems within the city.

We expect building information and crowd-sourced data to play a significant role in providing more data into the city environment.

8. Conclusions

There are a number of data trends that appear within smart cities and, most significantly, those that are required to make the city work for the citizen and communities that reside within the city.

Data needs to be made available allowing informed decisions to be made that provide support to the wide range and variation of communities to meet their specific needs. This need to serve a variety of communities means it is difficult to ignore any datasets as in a specific context any dataset can add value.

If you take for example the case studies in Glasgow, some use just five datasets. Others use more than twenty to achieve the desired outcome. Some of these datasets will be static, e.g. performance results for a school and others dynamic, like air quality and traffic flow.

The projects and initiatives within cities are also disparate. Taking London as an example, the Transport for London infrastructure shows a very high level of integration between public services providing a 'smart' travel experience. Similarly Barcelona is in a mature leading role capitalizing on sensor technology and network infrastructure to integrate its transport network. Barcelona also has a very mature framework; its "Anatomy Model" which structures the elements of the city into clear groupings. Glasgow city provides a similar approach to Barcelona, investing in sensor technologies but at the same time there is very clear evidence in understanding the core issues in the community and how best these are addressed.

Glasgow also quotes "90% of the world's data was created in the last 2 years" and the volume of data continues to grow at an incredible rate; so whilst datasets are today in the hundreds for any one city, it will not be long before this falls into the thousands.

So what will 'core datasets' actually mean?

Getting the private sector to open up its datasets for specific construction projects is common, but making these datasets freely available in support of wider smart city initiatives appears to be more of a challenge. For example, Utilities, Telco's, Civil Engineering companies, financial institutions and retailers all generate data that has a role to play in the development, growth and maintenance of the smart city. Data sharing will become a necessity as communities will demand more from their city. Smart city leaders and those responsible in the private sector will need to increase the volume and diversity of data sharing to address and resolve common problems.

BIM is clearly linked to smart city development and the approach to construction and the impact on the city and vice versa is a major consideration. Data that supports both topics will need to be shared, most likely under commercial arrangements as the whole life-cycle of a build and not just its construction needs to be considered in a sustainable city environment. The approach, leadership and ability to work with a shared vision will be critical for the success of their coexistence.

So where does this leave the subject of data and standards? Firstly, all cities expect to work primarily with open data. To maximize the open data opportunity there is a need to work more closely and commercially with the private sector to understand the potential of the data they hold, how data can be de-sensitized and how it can be made open. The benefits over time of smart city development will be enhanced by the ability to re-use data effectively and ensure interoperability of available data.

When posing the question 'What datasets do you consider essential for developing a smart city environment?' There is not one answer and as is evident in the Phase 2 survey results; it depends on the challenge and supporting projects needed to solve them.

What can be said though is that ideally any data:

- must be open;
- must be widely published;
- should be easy to ingest, although specific format may be less of a problem;
- should have content that is standardized for global interoperability.

From the results of the survey, these suggest that standardization should target datasets that address problems that will address:

- social and community;
- transport and mobility;
- infrastructure;
- technology.

Extending these four categories may merit standardizing data to include:

- emissions and improvements in air quality;
- smarter use of energy resources;
- improved city mobility and transportation;
- repurposing existing buildings to extend their life and use;
- improving the home environment;
- improving health and health management;
- efficiency improvements in national and international travel;
- helping communities to take ownership for decisions in the city.

It is fair to say that some of these areas are already bound by global agreements but may benefit from application of more detailed set of standards.

From the questions raised with the participants, Table 3 illustrates the key findings.

Key survey findings	
Who benefits the most from resolving the challenges within cities?	1. Residents 2. Economy 3. Local government
Proportion of open data being used:	1. Data you own (collected) **47%** 2. Open (free) data **31%** 3. Commercial (paid) data **22%**
Current largest sources of city data:	1. Local government 2. GIS providers 3. Real-time sensors
Barriers to obtaining data:	1. Anonymity – Removal of personal information from datasets 2. Competition – Utilities releasing customer and performance information 3. High costs to obtain, e.g. Mobile phone operators data for footfall and user profile 4. Costs of technology – Creating costly projects using sensors 5. Silos – Obtaining information across government departments within councils
Proportion of the data, cities have or will be using from available sources:	1. 54% Government 2. 15% Sensor technology 3. 16% Crowd-sourced 4. 15% Private sector
Datasets cities would like to use more in the future:	1. Mobile phone operator data 2. Crowd-sourced data.
Which of the following challenges are also relevant to your city?	1. Mobility / Transportation 2. Traffic congestion 3. Business generation and energy conservation

Table 3 – Key survey findings

Throughout this study, we have seen several examples of challenges that BSI is aiming to address through developing a decision-making framework for sharing data. These includes examples of the difficulties in sharing data due to departmental and commercial barriers, concerns over anonymity as well as identifying specific datasets to support key city initiatives.

Adopting a standard framework for acquiring, sharing and publishing data will not only aid interoperability but also aid the ability to apply analytics to disparate datasets in a consistent way allowing the improvement of city services, as well as like-for-like comparisons across cities nationally and internationally.

Appendix A – References and case studies

In examining the links provided by the interviewees, there are a large number of examples / case studies that are relevant to the smart city environment. It would be unhelpful to reproduce all of these here but for completeness we have chosen to reference some of those relating to Glasgow and Barcelona, Leeds and Stockholm by way of example. This provides a good spectrum of the types of work actively going on in these cities.

London

Name	Description	Hyperlink
London Air Quality Network	Pollution monitoring data feeds	http://www.londonair.org.uk/LondonAir/Default.aspx
London Living Labs	City-scale experimental projects	http://www.bartlett.ucl.ac.uk/energy/news/icri-cities

Leeds

Name	Description	Hyperlink
City Dashboard	Integrated display of city data	http://citydashboard.org/leeds/
Leeds Data Mill	A hub for citizens covering data, projects, education community and events	http://www.leedsdatamill.org/
Check your bin day	Simple tool to identify bin collection dates against a postcode	http://www.leeds.gov.uk/residents/Pages/Check-your-bin-day.aspx
Leeds Artcrawl	Facility to upload, map and share public art	http://leeds.artcrawl.club/welcome
Leeds City Council Air Pollution Monitoring	Live and historic data feeds	http://www.airviro.smhi.se/leeds/index.html
DC4	New data centre	https://www.aql.com/news/235/

Glasgow

Name	Description	Hyperlink
Glasgow Future City	The site established following winning of funding for the Future Cities Catapult (FCC) in a contest run by Innovate UK. It is pulling together the collaboration between public and private sector agencies providing a range of services to the city.	http://futurecity.glasgow.gov.uk/
Glasgow Data Launchpad	Open data portal	https://data.glasgow.gov.uk/
Glasgow Cycling	Cycling application (see case study in A.1)	http://futurecity.glasgow.gov.uk/index.aspx?articleid=14384 https://itunes.apple.com/gb/app/glasgow-cycling/id930422838?mt=8
Sensors in homes	Residents of Sandyhills have had sensors fitted in their homes to measure the benefits of different types of insulation	http://futurecity.glasgow.gov.uk/index.aspx?articleid=14074
OPEN Glasgow Community Map	Community-centric information platform	http://open.glasgow.gov.uk/datastories/mapping-by-the-community/
Intelligent street lights	Reactive LED lighting trial	http://futurecity.glasgow.gov.uk/index.aspx?articleid=10253

Peterborough

Name	Description	Hyperlink
Peterborough DNA	A resource for developing and testing ideas for a smarter city – funded as part of FCC.	http://www.peterboroughdna.com/
School Weather project	A network of school weather stations	http://www.ukauthority.com/news/5072/peterborough-school-weather-stations-hold-ict-front-line

Barcelona

Name	Description	Hyperlink
BCN Smart City	Barcelona's smart city web portal (see case study in A.3)	http://smartcity.bcn.cat/en
Barcelona Smart Trash Cans	Smart trash cans with sensors and wireless links that remotely alert cleaners when they are full	http://www.bloomberg.com/news/articles/2014-02-23/barcelona-s-smart-trash-cans-pave-way-for-mobile-future

Stockholm

Name	Description	Hyperlink
City of Stockholm the smart city	Stockholm smart city web portal	http://international.stockholm.se/city-development/the-smart-city
Stockholm city statistics	A comprehensive website dedicated to providing statistical information in many useful formats	http://www.statistikomstockholm.se
Open data Stockholm	A site dedicated to providing information on open data by category	http://open.stockholm.se/oppna-data
Stockholm's environment	Detailed facts on all aspects of the environment in Stockholm	http://miljobarometern.stockholm.se
Find and compare service	Detailed web site for citizens to get in contacts with a diverse range of services	http://www.stockholm.se/jamfor

Other UK references

Name	Description	Hyperlink
LG Inform	Presents an up-to-date published data about your local area and the performance of your council or fire and rescue authority	http://lginform.local.gov.uk/
Distance exploratory	A project which aims to bring together a series of Internet connected objects to enhance the classroom experience	http://iotschools.org.uk/DistanceExploratory/index.php

Other global references

Name	Description	Hyperlink
Socrata	Cloud solutions for open data and data-driven governments	http://www.socrata.com/
City protocol	a global non-profit community of cities, corporations, academic and non-profit organizations taking collaborative action to help cities face their challenges	http://cityprotocol.org/

Glasgow

For headline links to the Glasgow future city developments please see: **http://futurecity.glasgow.gov.uk/**

Grow Your Business
http://open.glasgow.gov.uk/datastories/grow-your-business/

Mapping by the Community, for the Community
http://futurecity.glasgow.gov.uk/community-mapping/

Leeds

City artcrawl
http://leedsdatamill.org/dataset/leeds-art-crawl

City dashboard
http://leedsdatamill.org/dashboard/

Barcelona

Illustration of smart city projects
http://smartcity.bcn.cat/en/apps4bcn-portal.html

Stockholm

Stockholm is particularly advanced in developing solutions across the whole range of issues experienced in a smart city environment. This ranges from Green IT strategies through to traffic management, elderly care regimes and many others. Key links are provided below:

http://international.stockholm.se/city-development/the-smart-city/

http://www.statistikomstockholm.se/

http://open.stockholm.se/oppna-data

http://miljobarometern.stockholm.se/

http://www.stockholm.se/jamfor

Appendix B – Survey questions and results

B.1 Phase 1 survey question

1. What datasets do you consider essential for developing a smart city environment?
2. Where do you expect to obtain these datasets – open source, internal or other?
3. Which datasets are you already using and for what purposes?
4. What supporting case studies that exploit these datasets can you share with us?
5. In the context of smart cities, what data do you currently or expect to derive and publish to others to contribute to making your city 'smart'?
6. Who else in other smart city initiatives do you liaise with that you would recommend we contact to further progress this research?
7. What hyperlinks are you willing to send us to allow us to access any useful resources that will expand or support the conversation? For example, case studies.

B.2 Phase 1 results

B.2.1 Participating cities

Country	City represented
Spain	Barcelona
Great Britain	Bristol
Emirates	Dubai
Great Britain	Glasgow
Singapore	Singapore
Great Britain	London
Great Britain	Greenwich
Great Britain	Leeds
Great Britain	Peterborough
Great Britain	Sedgemoor
China	Beijing (MoHurd)

Because Phase 1 was conducted as telephone interviews we have limited the response to the questions to the understanding of the datasets (the focus of the survey) each city is or expects to use as part of their city development. These results are shown in tables B.1 to B.3.

Data set requirements by city category (phase 1)

Barcelona	Building	Emissions Supply Demand Maintenance	Supply Demand Recycling Natural flow Maintenance	Garbage collection – bin sensors Mobility		Climate and weather Green areas Service improvement Bio-diversity Geographic features	Cell data Fiber networks People density	Data from Crowd sourcing		Fixed sensors Mobile sensors Control
Bristol		Supply Demand Maintenance			Traffic flow Terrain / topology	Air Quality Tree locations Weather conditions Land Use		Data from Crowd sourcing Crime Fuel Poverty	Health statistics	Fixed sensors
China (MoHURD)	Geographic GIS data Building environment BIM / Physical data							Demographic data about the citizen, aging, education		
Singapore	Built environment for planning purposes		Flood and drainage		Transport Moving people from A to B	Weather API's published		Demographics, aging population, where people live and work placement of services to support them		
Dubai		Electricity bills						Living standards Housing Information		
GLA	Planning Infrastructure maintenance schedules	Usage statistics Energy ratios for properties	Usage statistics Supply Demand	Transport data	Vehicle recognition data		Mobile data but expensive!	People movement Crowd source		

Glasgow	Asset location data	Usage statistics Street lighting data	Usage statistics Supply Demand	All transport types Services optimisation	Cycle data Traffic flow	Land Use		Community data sets Demographics Crime data Public safety data Fly tipping data	Pest data	Sensors data
Greenwich	GIS Backdrops Building density and height over 6m	Shared heating services from power station			Driverless cars		High speed fiber	Child services Aging population Census data		
Leeds	Business rates Industry data Retail data	Home Insulation Stats.	Supply Demand Maintenance	Garbage collection	Geographic location Traffic flow	Pollution / Air quality Traffic flow (pollution) Hazard data e.g. trees Met Office (Weather)				
Peterborough	Planning Improvements to buildings resulting in improved health lifestyle				Feature identification to improve mobility	Air quality General weather data		Crime Population	Health statistics	
Sedgemoor District Council	Data to share planning information			Data to identify service org.		Air Quality		Data that defines communities Data to inform political agenda Data to identify service org. Data to share planning information Data to inform funding decisions	Data on diseases	
Autodesk Eng.	3D data sets Data to aid repurposing / refurbishment of properties	Utilities physical location Emissions	Supply and drainage physical location	Transport data	Traffic information	Air Quality				

B.3 Phase 2 survey questions

The survey presented each participant with 22 questions which were a mix of multiple choice and free text fields. These are listed below:

1. Thinking about the current challenges your city is facing, what do you consider as the biggest challenge at the moment?
2. Thinking about this specific challenge... How beneficial would it be to the following if you were able to improve its current situation?
3. Thinking about this specific challenge... What projects are you working on to tackle it?
4. Please tell us the top three critical datasets that you need for these projects?
5. Now thinking about the current challenges your city is facing, what do you consider as the second biggest challenge at the moment?
6. Thinking about the second biggest challenge... How beneficial would it be to the following if you were able to improve its current situation?
7. Thinking about the second biggest challenge... What projects are you working on to tackle it?
8. Please tell us the top three critical datasets that you need for these projects?
9. Lastly, what do you consider as the third biggest challenge your city is facing at the moment?
10. Thinking about the third biggest challenge... How beneficial would it be to the following if you were able to improve its current situation?
11. Thinking about the third biggest challenge... What projects are you working on to tackle it?
12. Please tell us the top three critical datasets that you need for these projects?
13. The following questions will refer to your overall data requirement... Thinking about the data required in general, which of the following did you use or will be using? (Please tick all that apply.)
14. What is the proportion of open (free) data versus paid data you used in general?
15. Do you have any budget for purchasing data?
16. What is your typical budget for purchasing data, rounded into US dollars:
17. Which of the following sources have you used to obtain the data required? (Please tick all that apply.)
18. On a scale of 1-5 where 1 is very difficult and 5 is very easy, how would you rate the easiness of getting data from the following sources?
19. Can you please tell us which data sources are most difficult to get data from and why?
20. Can you also use the table below to let us know what proportion of the data you have used / will be using is from the following sources?
21. Which of the following data sources that you are not using but would like to use in the future? (Please tick all that apply.)
22. Apart from the top challenges you mentioned in the previous question, which of the following challenges also relevant to your city? (Please tick all that apply.)

B.4 Phase 2 survey results

In Phase 2 we adopted the use of an online survey approach to gather information from a range of cities throughout the UK, Europe and further afield. The cities that supported the survey are as follows.

B.4.1 Participating cities

Country	City represented
Scotland	Aberdeen
Great Britain	Birmingham
France	Bordeaux
Great Britain	Cardiff
France	Cerema organisation
Denmark	Copenhagen
Spain	Coruña
Belgium	Ghent
Spain	Malaga
Israel	Ministry of National Infrastructures, Energy and Water
Germany	Munich
Great Britain	Perth
Australia	Perth
Great Britain	Preston
Croatia	Rijeka
Spain	Santander
Sweden	Stockholm
Finland	Tampere
Spain	Valencia
Austria	Vienna

The responses to the 22 questions were compiled by the survey provider generating the following results:

Questions 1 to 12 asked about the cities top three challenges they were addressing or aiming to address, the projects they were running or initiating to address these challenges and the datasets that would be needed to support these.

B.4.2 Phase 2 responses

Challenges, projects and data set requirements

City	No.	Challenge	Project (list)	Data Set 1	Data Set 2	Data Set 3
Aberdeen	1	Being addressed as part of Scottish Cities Alliances digital programme inter alia	Many. These range from digital infrastructure investigation, co-operation across Scotland on smart cities and innovation programmes to hack weekends and the broader use of open data. Often these have specific themes such as transport, environment, health etc.	Yet to be determined	Yet to be determined	Yet to be determined
	2	Digital literacy within the organisation and externally	Designing internal training, and working with external partners to embed skills, behaviours and knowledge in partners, citizens, and others	Yet to be determined	Yet to be determined	Yet to be determined
	3	Transport and supporting infrastructure	Being addressed as part of Scottish Cities Alliance digital programme inter alia.	Yet to be determined	Yet to be determined	Yet to be determined
Gent	1	Mobility, sustainability, community management				
Preston	1	Establishing the most appropriate governance, resources and policy package at the city level so as to best promote and deliver economic and population growth.	Working together with other medium sized cities to make the case for greater devolution (Key Cities). Working with other authorities to establish a Combined Authority for Lancashire. Establishing a common Local Plan for our Functional Economic Area (common planning, economic development, housing & transport plan covering four authorities). Delivering an infrastructure based City Deal across the contiguous urban area of the city. Working with other local "anchor institutions" to reduce "leakage" of public procurement spend out of the immediate area and build a "good leak economy"based on principles of economic democracy and social justice.	GVA at level of functional economy (not LADS or Nuts2	Travel to work statistics on FEA basis	More robust data sets on well-being
	2	Addressing the barriers to higher levels of economic participation	Promoting the Living Wages; working on deliver of a local "Fairness Charter" across public, private and third sector partners; a range of activities designed to remove barriers to participation in the labour marker by disadvantaged communities and enhance the skill base of the local population, including targeted training & recruitment for jbs created through the City Deal.	Measurement of economic participation rates relative to total population	Level and quality of jobs available, including rates of pay	Reduction in cost to state in responding to poverty/ social exclusion
	3	Responding in an informed and well-judged way to the new demands and opportunities offered by new technologies (ICT) within the context of constrained public finances.	Super-fast City Broadband Voucher Scheme. Work with other public and private sector partners on enhancing city broadband facilities. Renewing and re-imagining the city council's own ICT systems. Promoting ICT Skills development and access to ICT through community-based projects etc.	Reliable & comprehensive data to ward level on household & business access to and take up of high speed broadband	Robust survey information on ICT specific skills in local population	Up to date information capacity of broadband infrastructure locally

City	No.	Challenge	Project (list)	Data Set 1	Data Set 2	Data Set 3
Vienna (1)	1	Growing while saving energy	Resource preservation development and productive use of new technologies high and socially balanced quality of living	High quality data sets about resource usage	High quality data sets about all city geometries	High quality demographic data sets
	2	Keeping a high degree of social cohesion and quality of living	Social housing reduction in CO2 emissions	High quality data about social mobility	High quality data about labour mobility	High quality demographic data sets
	3	Financing long lasting infrastructure	Looking for economic models difficult, because of rising debts	Change of infrastructure in cities	Basic economic data on possible fields of actions	
Vienna (2)	1	Change of the economic game worldwide – find a position in this new game	Force local economy in hi-tech branches Finance research education, universities, etc.	Economic development in branches	Innovation related education data	
	2	Climate change	Energy efficiency in housing and mobility sector Local use of renewables and waste heat Force local economy in the field of climate change technologies	Energy use of the housing sector	Mobility (modal spilt etc.)	Greenhouse gas emissions
Valencia	1	Economic situation – levels of unemployment	Helping economics evolution to a knowledge and innovation on economy. Fostering innovation ecosystem, open data for reusing.	Activities, commerce and industry	Real-time data on traffic and transport	Innovation resources
	2	Environmental challenges	Smart city platform, improving energy efficiency, reducing carbon emissions	Pollution	Traffic and transport	Energy consumption
	3	Social issues, ageing population, health system, accessibility, inclusive society	Technology for elderly people, universal accessibility,	Population	Health information, ex defibrillator	Disabled parking
Tampere	1	Rigorous economic situation	Several digitization projects	Customer Data	Service/ product data	Financial data
	2	Supporting new business challenges	Open data projects, smart city projects	Traffic data	Spatial data (i.e. maps)	Urban planning and buildings
	3	Re-organizing welfare services	New digital services	Customer data	Queuing data/ utilization	Service/ product
Copenhagen	1	Congestion, cloudburst/torrential rainfall, Air quality issues, Silo thinking	Climate plan and climate adaptation plan, and smart city strategy	Real time traffic data	Sensor data from parking	Air quality data
	2	Silo thinking	Holistic thinking and labs working cross border	New generations	New competencies	Smart city thinking
	3	To implement smart city projects	Copenhagen connecting, a holistic silo breaking IoT approach www.cphsolutionslab.dk	Traffic data	Parking	Waste

B.4.2 Phase 2 responses – Challenges, projects and data set requirements continued...

City	No.	Challenge	Project (list)	Data Set 1	Data Set 2	Data Set 3
Rijeka	1	Developing a competitive economy on the basis of the knowledge society and the new technology	Support for developing "new" i.e. creative industries Support for transfer of knowledge and technology from university to industry	Utility infrastructure (underground and aboveground)	The current position of the bus public transport in time	Commercial property of the city
	2	Contemporary social policy Support of Health lifestyles (prevention programs, sports, recreation)	Contemporary social policy Support of Health lifestyles (prevention programs, sports, recreation)	Social policy measures	Locations and events of recreation and sport	
	3	Investments in the infrastructure Global Promotion and international cooperation Strengthening dusters in logistics and maritime affairs	Investments in infrastructure Global promotion and international cooperation Strengthening dusters in logistics and maritime affairs	Cadastre	Traffic data	Cargo data
Birmingham	1	Youth skills and unemployment. Birmingham has a large you population, the skills need so reflect the industries requirements for now and the future. Many deprived areas in the City are not meeting this challenge.	City demonstrator projects to help promote jobs and skills in the Eastern Corridor area. Looking it build towards the impact of HS2 and the HS2 college that will be built to grow skills in Infrastructure. IT processes working with he DWP to match small business to skills and opportunities that young people can provide or benefit from.	Small business data	Existing skill sets	Local jobs
	2	Social Care for Adults and vulnerable children	Partnership working between academic institutions, public health, the NHS, police, child safeguarding and other institutions. Pulling data together to integrate and understand the pathway that a person takes in their interactions wit these institutions to identify problems and move to more predictive modelling	Young people who are "Not in Education, Employment, or Training" NEETs	Uptake of social care packages (in private sector)	Movement of pupils between academic institutions
	3	Creating the environment for economic regeneration and growth	Understanding business data, economic output, locations of business, the types of businesses and how the supply chain works. Access to work for mobility purposes and crucially the institutions to support this – the Local Enterprise partnership and the Combined Authority (like Greater Manchester has) to enable confidence for devolution	Data	Sector growth and supply chains (flow of money)	Detailed journey to work data (incorporating school drop offs as leads to huge congestion)

City	No.	Challenge	Project (list)	Data Set 1	Data Set 2	Data Set 3
Malaga	1	Coordination of the different systems of the city	Unified Control Center, Inegration of call centers, development of a city platform	Not Known	Not Known	Not Known
	2	Improve mobility and make it more ecological	Project about electrical vehicle, cycle paths, electrical buses	Not Known	Not Known	Not Known
	3	Employment	Improving the tourism (create jobs), enhancing technological park, making accelerators and incubators	Not Known	Not Known	Not Known
Cardiff	1	The future (50 years) and the infrastructure needs and how smart and integrated they can be	Working with the likes of Siemens and RS, Cardiff Uni and BRE to secure TSB or European funding to pilot smart infrastructure approaches	Energy	Transport	Waste/ Water
	2	Getting other people to also accept this as the challenge	We don't have space, budget or capacity to do this			
	3	We are due to build 40,000 new homes by 2026 in the city and our developers (wimpey, Permission, Redrom) are a million miles away from this agenda	Asking if Welsh Government can help us bring these stakeholders together	Performance of smarter buildings	Construction costs	Customer response
Coruna	1	Improvement and sustainability of public services	Smart Coruña Programme. This is a Smart City project and several pilots aimed to improve the management of public services	Assets inventories	Real time information	Quality indicators
	2	Improvement of the citizens quality of life	Smart Coruña Programme			
	3	Improvement of local economy	Smart Coruña Programme			
Cerema (France)	1	To supply a large offer of efficient customizable services based on pre-treated and real-time delivered by very few but familiar front-end smart phone applications (applies mainly to transportation).	The most representative example is ticketing (in French "billetique") (see AFIMB agency of French transport ministry). The goal is to have a standard along the chain data-equipment-services in order on the one hand to ease the system management for the local authorities and the other to deliver multi-service to users (OURA card is the key example in the Rhône-Alpes Region). Cerema is one of the expert pool on this subject. Another Example is managed with "grand Lyon" metropole with Optimod as a predictive service for travel optimization. In addition, Cerema is also working on sata standardization (COVADIS commission for example is one equivalent of BSI and is working on thematic standards applying to state services as well as local authorities).	Open networks cartography	Traffic schedules	Big data (from GSM and other connected devices) to know the traffic, the speed...
	2	Comprehensive and global knowledge about Return Of Investment (economics, efficiency and so on).	Just starting benchmarking, and indicator definitions and observation			

B.4.2 Phase 2 responses – Challenges, projects and data set requirements continued...

City	No.	Challenge	Project (list)	Data Set 1	Data Set 2	Data Set 3
Perth	1	The biggest challenge is balancing a softening economy and business conditions with a range of major new developments that are under construction and about to be finalised.	The City is placing considerable resources into ensuring the new developments are normalised within the city and have successful and enjoyable public realms. The City is also working on the diversification of the local economy and encouraging growth of alternate sectors to make Perth's economy less reliant on the resources sector	Office market / vacancy rates	Daily visitor figures	Population figures and projections
	2	Trying to transition into a "smarter city" and capturing and utilising all the data that this entails	Current organisational restructure will see the creation of a new "Data and Information" Unit to maximise the opportunities can take.	Spatial data	Land use	Development information
	3	Dealing with issues associated with strong population growth	Advocating for improved public transport infrastructure and encouraging diversity of housing options	Population figures and forecasts	Transport patronage and forecasts	Housing mix
Bordeaux	1	The biggest challenges for Bordeaux Urban mobility	Contribution to European H2020 MG calls: Compass 4D, Smartline, Intramo, etc... A proactive approach of transport public open data	Real-time parking availability	Real-time public transport information	Real-time citizen mobility
	2	Urban sustainable refurbishment. Low energy districts conception	A territorial climate and energy plan. A broadband network roll up for every Citizens. Contribution to European H2020 SCC calls, deployment of proof of concepts around urban lighting, Building Energy Management Systems...	Utility counters vs thermography	Real-time occupancy of buildings	Real-time production capacity in renewables of a district
	3	Citizens empowerment about their environment	A territorial climate and energy plan. Animation of 122 families to adopt eco-citizens behavior. Crowd-sourcing applications: "Bordeaux proximité".	Reliable and comprehensive data toward level on household & business access to and take up of high speed broadband	Robust survey information on ICT specific skills in local population	Up to date information on capacity of broadband infrastructure locally

For each challenge identified by the city they were asked to consider the benefits on a scale of 1 to 5 for:

- local government;
- central government;
- residents;
- industry;
- economy;
- environment.

Four cities did not cite any benefit alongside their projects and some 'did not know' specific benefits for certain projects. The information from individual cities identified that the challenges they were addressing would provide good benefit across all groups but central government and industry may not benefit quite as much. Overall though the balance of benefits would suggest that the challenges being addressed at a strategic level are serving the needs of most groups.

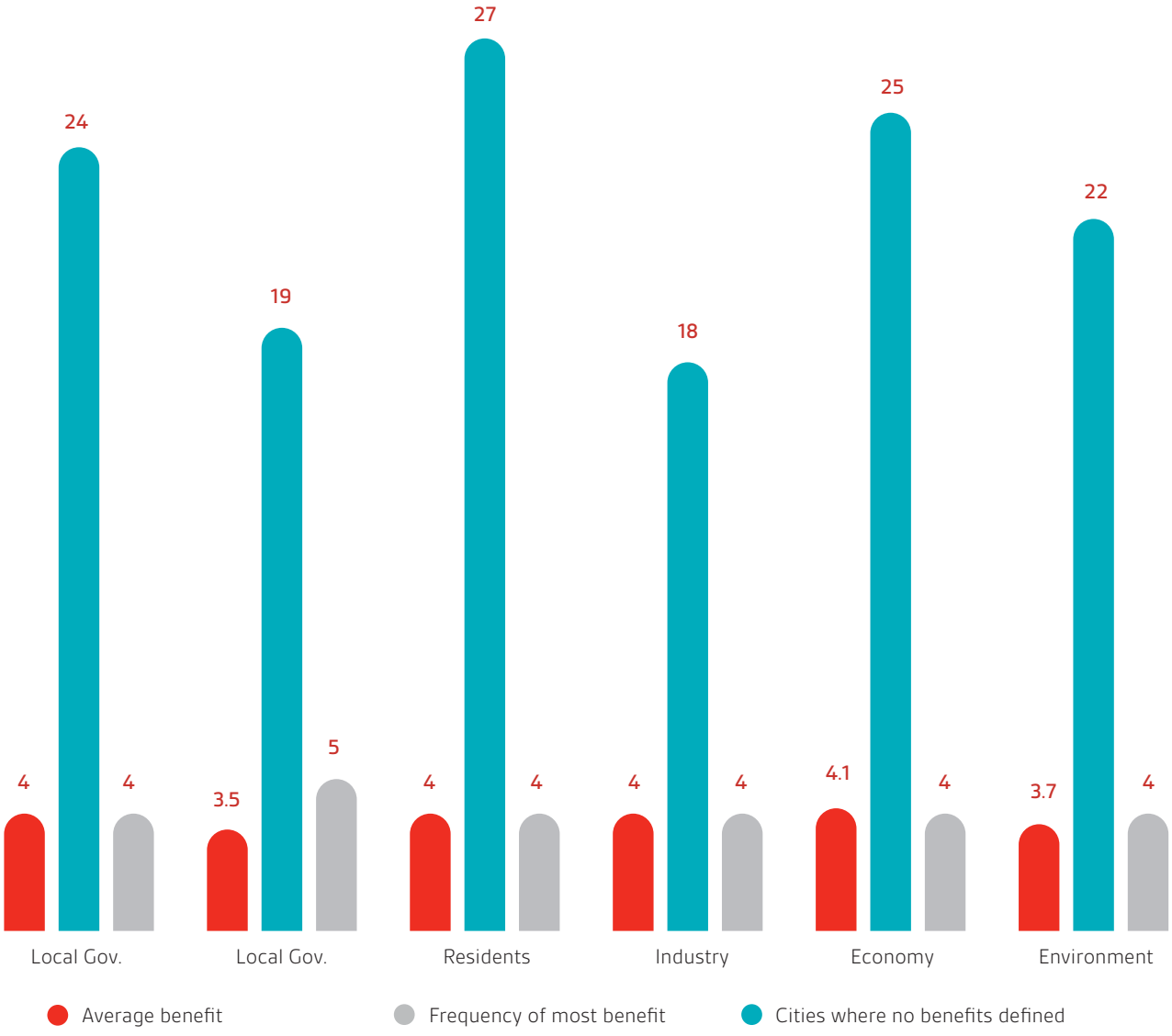

Figure B.1 – Relative benefits

13. **The following questions will refer to your overall data requirement... Thinking about the data required in general, which of the following did you use or will be using? (Please tick all that apply.)**

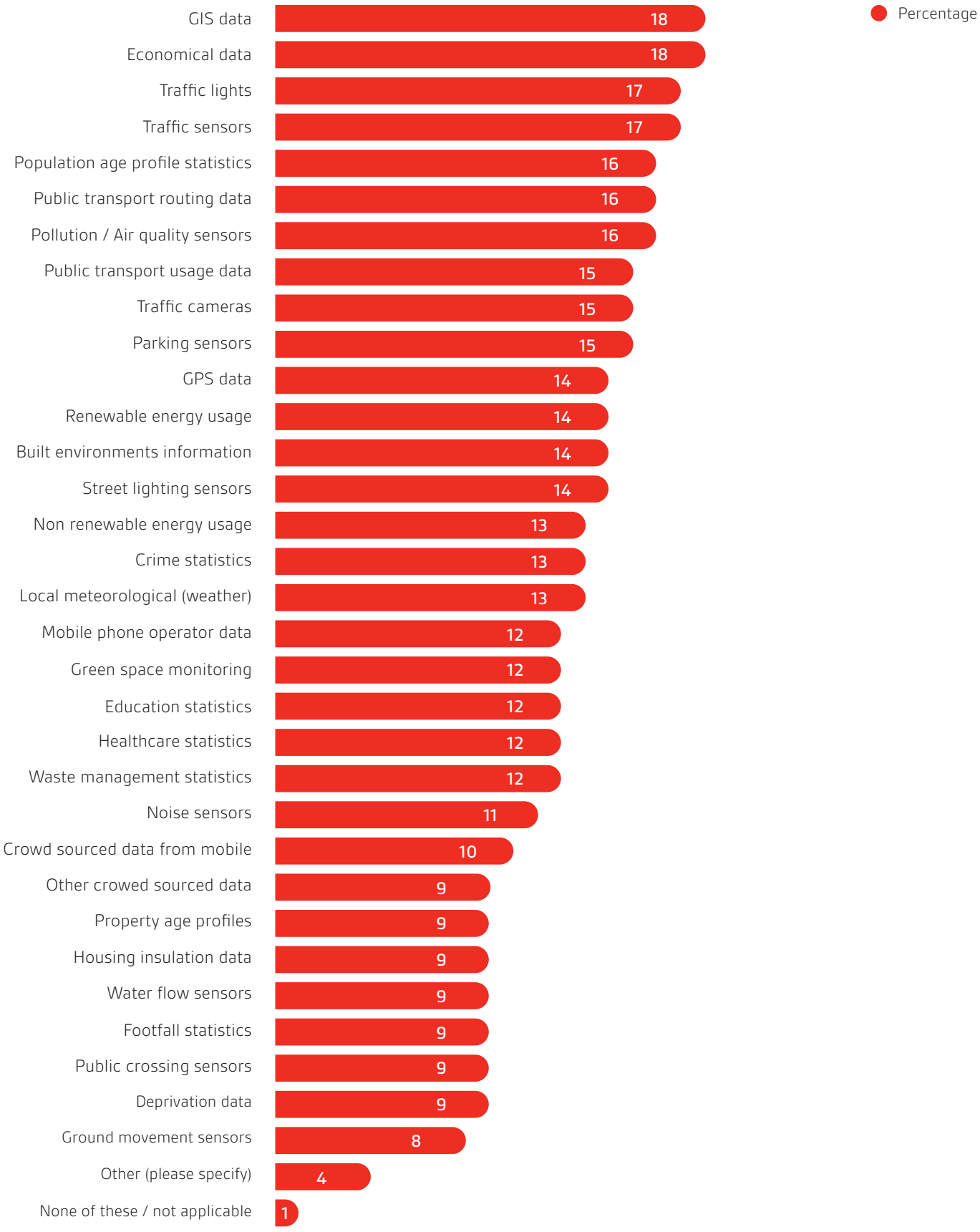

Figure B.2 – Occurrence of datasets being used or planned to be used

14. What is the proportion of open (free) data versus paid data you used in general?

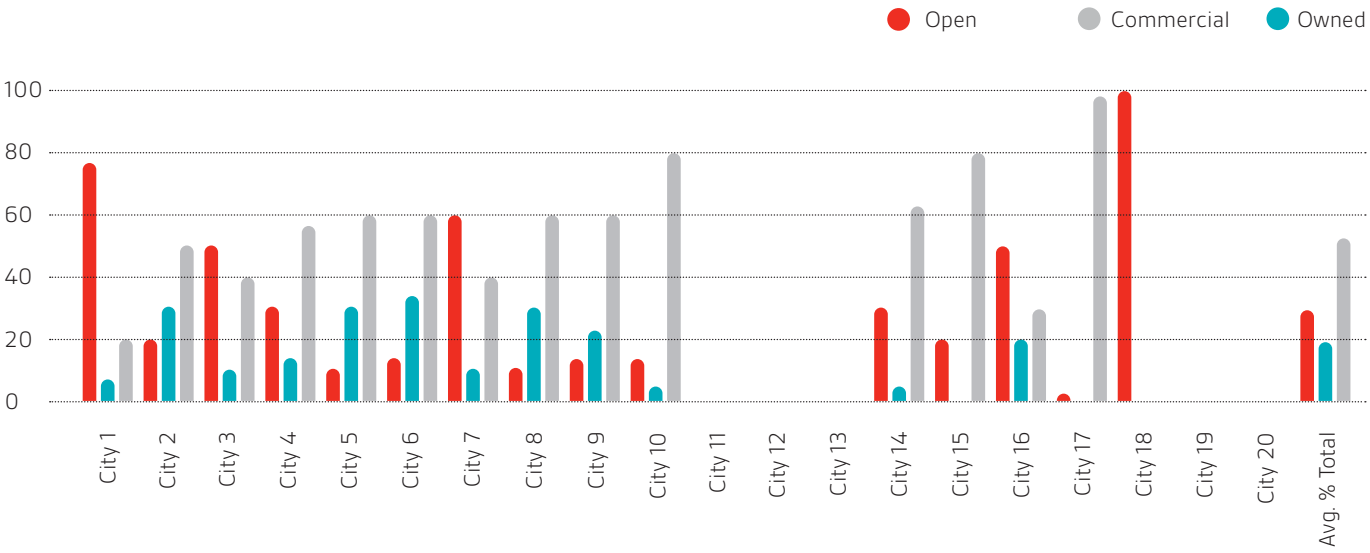

Figure B.3 – Proportion of open (free) data, commercial and owned data

On average, owned data was dominant with open data second. Data from Industry where this had to be paid for was consistent with Phase 1 in being third.

15. Do you have any budget for purchasing data?

Only two cities chose to respond confirming they have budget.

16. What is your typical budget for purchasing data, rounded into US dollars?

Both cities who responded suggested they had a budget of $100,000.

17. Which of the following sources have you used to obtain the data required? (Please tick all that apply)

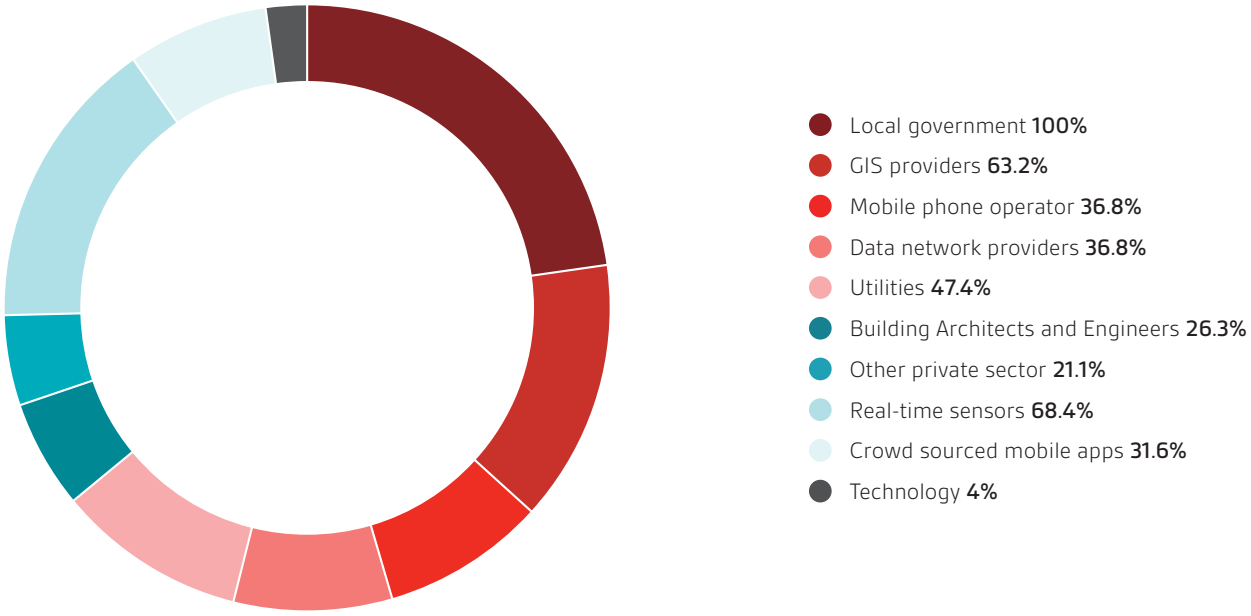

Figure B.4

	Percentage	Number
Local government	100%	19
GIS providers	63.2%	12
Mobile phone operator	36.8%	7
Data network providers	36.8%	7
Utilities	47.4%	9
Building architects and engineers	26.3%	5
Other private sector	21.1%	4
Real-time sensors	68.4%	13
Crowd-sourced mobile apps.	31.6%	6
Other (please specify)	10.5%	2
	Respondents	19
	No response	1

18. On a scale of 1-5 where 1 is very difficult and 5 is very easy, how would you rate the easiness of getting data from the following sources?

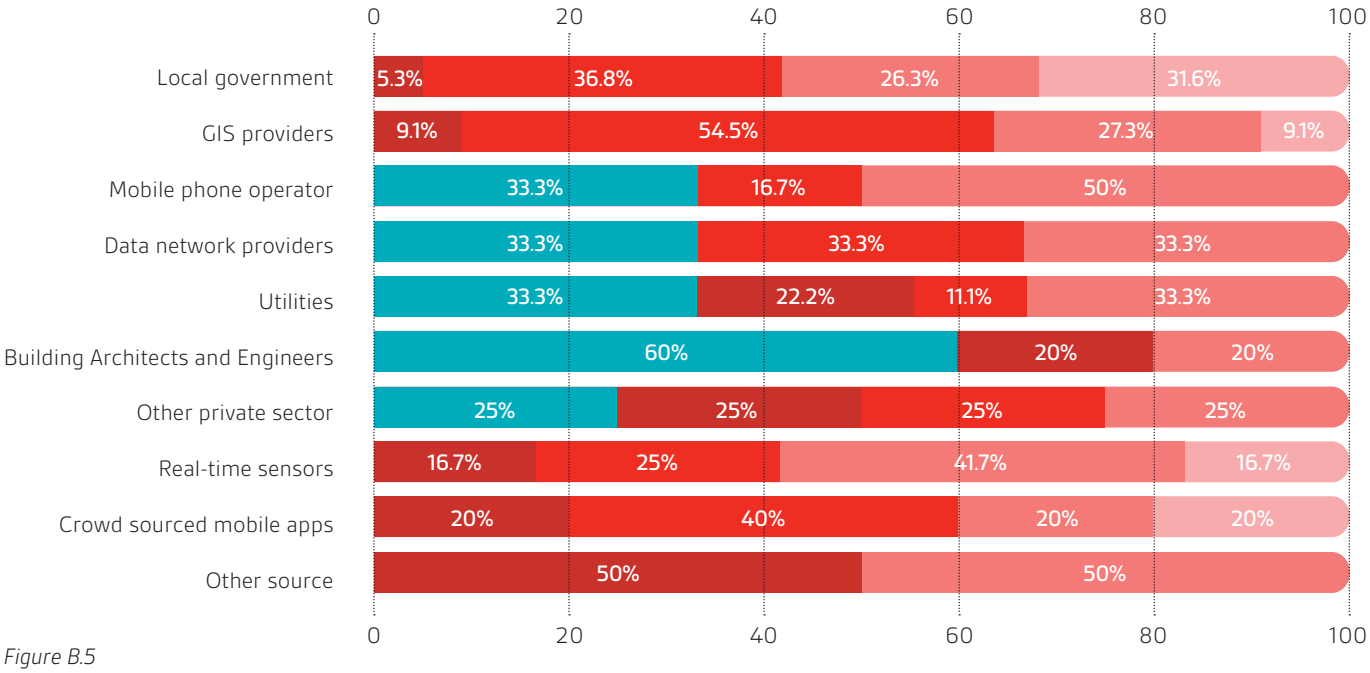

Figure B.5

	1	2	3	4	5	Respondents	No. of responses
Local government	0%	5.3%	36.8%	26.3%	31.6%	19	0
GIS providers	0%	9.1%	54.5%	27.3%	9.1%	11	1
Mobile phone operator	33.3%	0%	16.7%	50%	0%	6	1
Data network providers	33.3%	0%	33.3%	33.3%	0%	6	1
Utilities	33.3%	22.2%	11.1%	33.3%	0%	9	0
Building architects and engineers	60%	20%	0%	20%	0%	5	0
Other private sector	25%	25%	25%	25%	0%	4	0
Real-time sensors	0%	16.7%	25%	41.7%	16.7%	12	1
Crowd-sourced mobile apps	0%	20%	40%	20%	20%	5	1
Other source	0%	50%	0%	50%	0%	2	2
					Total	19	0

The percentage figures relate to the percentage of total respondents. For example, for GIS operators, 54.5% of the 20 respondents felt that getting GIS data was of average (3) difficulty.

From Figure B.5 we can see that by far the most difficult source of data is from Building architects and engineers, utilities and data network providers.

19. **Can you please tell us which data sources are most difficult to get data from and why?**

Local government	GIS providers	Mobile phone operator	Data network providers	Utilities	Other private sector	Real-time sensors
				Secret because of business		
NA			Networks topology, users properties, for business secrecy	Counters, they sell the datasets		
Data protection issues/ nervousness				Western power have been useful on all fronts		Half hour energy meters
		Lack of catalogues and/ or metadata, inappropriate cost, lack of standards, excessive use terms, black box effect	Lack of catalogues and/ or metadata, inappropriate cost, lack of standards, excessive use terms, black box effect			
					Consider data as their competitive advantage	
Anonymized personal data	Generally very good	High cost, questionable context data for footfall and traffic flow		Very guarded		Requires IT projects to release the data in a usable way

No comments were made relating to data from Building architects and engineers or for crowd-sourced data.

In summary, there are five areas to be considered when understanding the difficulty in obtaining information:

1. **Anonymity** – Removal of personal information from datasets.
2. **Competition** – Utilities releasing customer and performance information.
3. **High costs to obtain**, e.g. Mobile phone operators data for footfall and user profile.
4. **Costs of technology** – Creating costly projects using sensors.
5. **Silos** – Obtaining information across government departments within councils.

The last point did not come out directly from the Phase 2 survey but it was prevalent in Phase 1 through talking to city representatives.

20. Can you also use the table below to let us know what proportion of the data you have used / will be using is from the following sources?

Proportion of the data used from each source

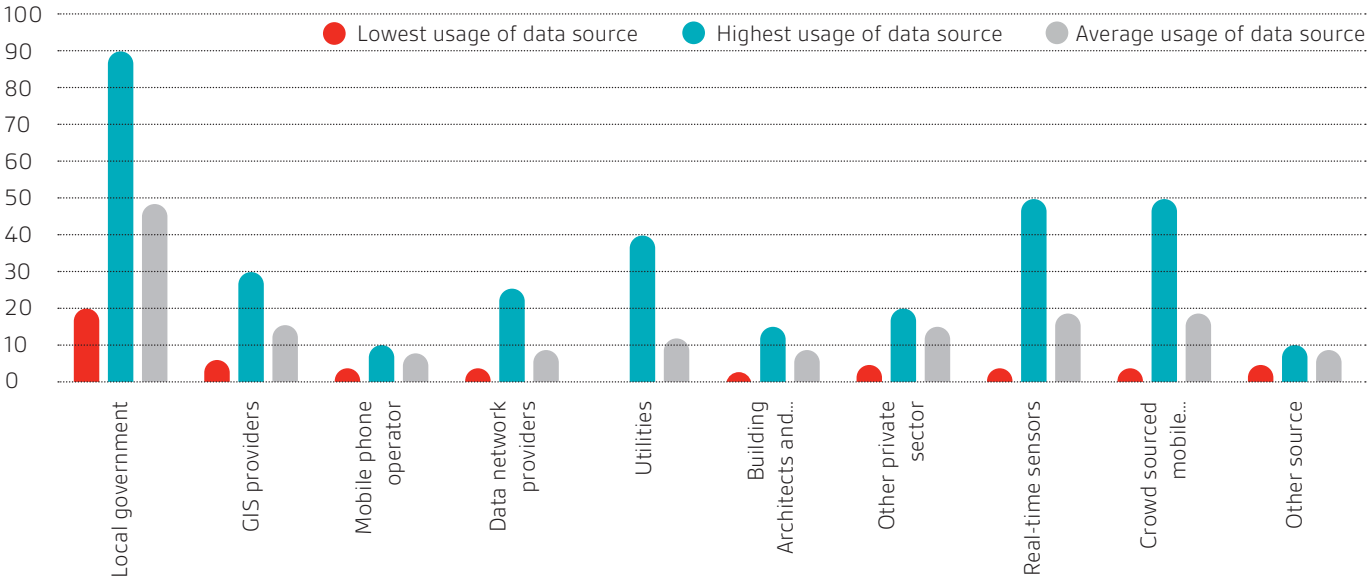

Figure B.6 – Proportion of the data used from each source

Most data that the city uses is provided by local government followed closely by the use of real-time sensors and utilities. Data from Building architects is not playing a significant role in smart cities and this is attributed to the reasons identified in question 19.

Proportion of the data used from each source by city

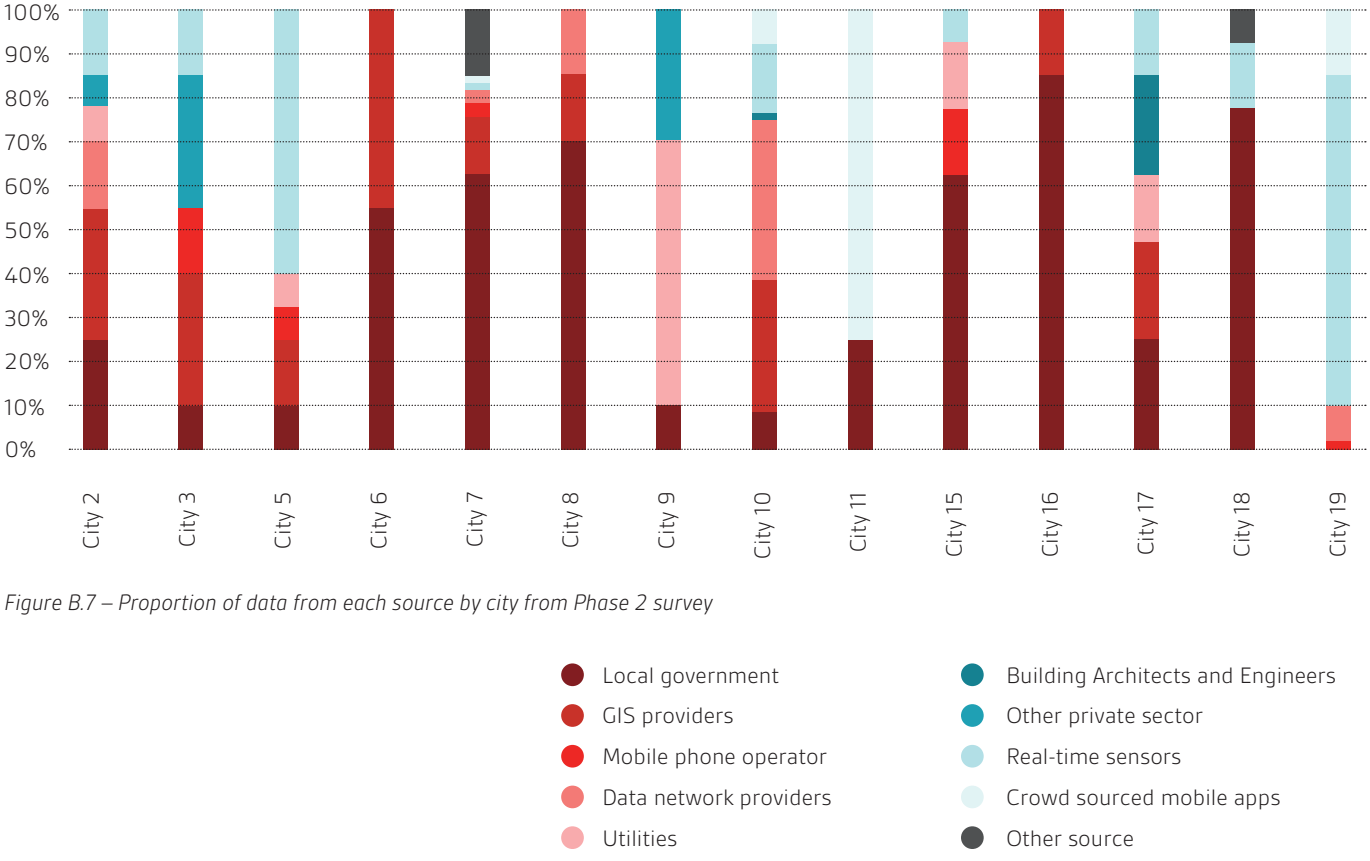

Figure B.7 – Proportion of data from each source by city from Phase 2 survey

21. Which of the following data sources that you are not using but would like to use in the future? (Please tick all that apply.)

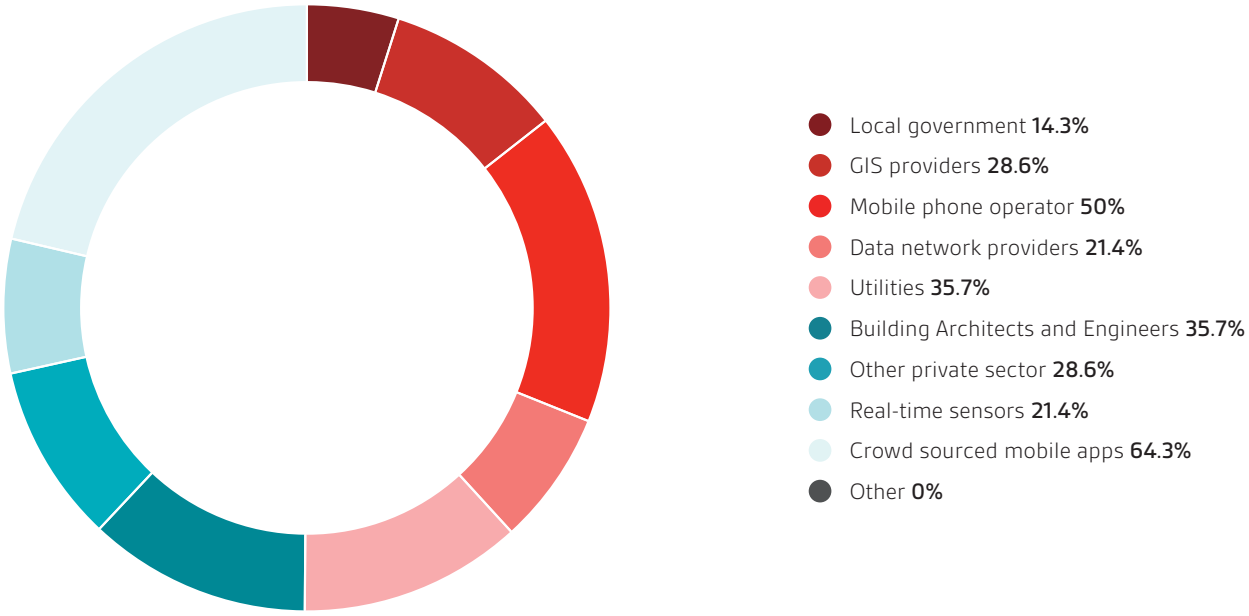

Figure B.8 – Data sources expected to be used in the future

	Percentage	Number
Local government	14.3%	2
GIS providers	28.6%	4
Mobile phone operator	50%	7
Data network providers	21.4%	3
Utilities	35.7%	5
Building architects and engineers	35.7%	5
Other private sector	28.6%	4
Real-time sensors	21.4%	3
Crowd-sourced mobile apps.	64.3%	9
Other (please specify)	0%	0
	Respondents	14
	No response	4

22. **Apart from the top challenges you mentioned in the previous question, which of the following challenges are also relevant to your city? (Please tick all that apply.)**

- Population segmentation **46.7%**
- Traffic congestion **80%**
- Mobility **86.7%**
- Healthcare provision **60%**
- Housing management **53.3%**
- Waste management **60%**
- Toxic waste management **20%**
- Energy conservation and renewables **73.3%**
- Environmental balance **46.7%**
- Security **46.7%**
- Pollution toxic emissions **33.3%**
- Climate management **46.7%**
- Cultural improvements **53.3%**
- Natural disasters (Flood, Earthquake, landslide, volcanic) **40%**
- Business generation **73.3%**
- Other (please specify) **0%**
- None of the above **0%**

Figure B.9 – Challenges within cities

Figure B.9 shows the occurrence of the challenges cities foresee in the future. **Mobility and traffic congestion rank highest with business generation and energy conservation coming equal second.** The chart shows that the challenges for cities are diverse but that there are common problems that will need to be addressed and will therefore need appropriate datasets to address these challenges.

BSI Group
389 Chiswick High Road
London, W4 4AL
United Kingdom

T: +44 20 8996 9001
bsigroup.com