

A Survey of Open Data Platforms in Six UK Smart City Initiatives

Journal:	<i>The Computer Journal</i>
Manuscript ID	COMPJ-2017-11-0717.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Ghahremanlou, Lida; Microsoft Research, Data and AI Special Team Unit Twavil, Abdel-Rahman; Birmingham City University Faculty of Technology Engineering and the Environment, Computing Kearney, Paul ; Birmingham City University Faculty of Technology Engineering and the Environment Zhao, Xia; Birmingham City University Faculty of Technology Engineering and the Environment Nevisi, Hossein ; Birmingham City University Faculty of Technology Engineering and the Environment Abdallah, Ali; Birmingham City University Faculty of Technology Engineering and the Environment
Key Words:	Open Data, Smart Cities, Innovative City Tools, City Applications



A Survey of Open Data Platforms in Six UK Smart City Initiatives

LIDA GHahremanlou^a, ABDEL-RAHMAN TAWIL^b, PAUL KEARNEY^b,
XIA ZHAO^b, HOSSEIN NEVISI^b, ALI ABDALLAH^b

^aMicrosoft, Thomas Valley Park, Reading, United Kingdom
lida.ghahremanlou@microsoft.com

^bBirmingham City University, Birmingham, United Kingdom

Email: {abdel-rahman.tawil,paul. Kearney,xia.zhao,hossein.nevesi,ali.abdallah}@bcu.ac.uk

This paper presents a comparative analysis of the feasibility studies submitted by six UK cities (London, Birmingham, Manchester, Glasgow, Bristol and Milton Keynes). The datasets were collected during the period of February to March 2017 by exploring their Open Data resources, common visions of their smart programmes and the themes of their projects. The aim of this work is to raise awareness and access of the existing data resources, encourage alignment of data collection and curation among projects with compatible objectives in different cities, and to identify the gaps in coverage that are hampering achievement of the cities' visions. In this research, we distinguish between stored datasets that are accessible via data hubs, and live data that are only accessible via APIs in real-time. Given that our findings are purely based on stored Open Data, we conclude that the Smart City dream will be only achieved in reality, where those involved in Smart Cities related projects co-operate and share both experiences and resources in order to maximise progress towards the common goal but to minimise duplication of efforts and repetition of the same mistakes.

Keywords: Open Data; Smart Cities; Innovative City Tools; City Applications

1. INTRODUCTION

Cities are vital to the future global economy. According to the United Nations reports [1], for the first time in human history, more people lived in urban areas than outside of them in 2008. By 2050 more than 70% of the global population will live in cities. In the UK cities are equally important, with one third of the country's total population living in the ten largest urban areas. This would result in adding 2.5 billion people to the world's city-dwelling population. The fact has significantly increased the attention of **Smart Cities** initiatives around the world, which results in the strategic response by governments to this growth.

Many cities throughout the world seek to improve the services provided to their citizens and to make their operations more efficient through the application of digital technologies, thereby achieving the status of Smart Cities. There appears to be little consensus, however, as to what qualities make a city smart, or what needs to be implemented to acquire them.

Public authorities are seeking to reduce costs and raise performance by adopting empowered Web-based approach in the delivery of quality public services. Accordingly, the concept of smart cities is getting increased attention in the media, from technology

companies and industrialists and endorsement in funding from both local governments and civil society. Yet, it is difficult to give an exact definition of a smart city due to the wide range of the technologies and applications that have been implemented under this label. The different definitions of smart cities in the related literature can be mainly addressed from three points of view as follows:

Broad definitions: The British Standards Institution (BSI) defines a Smart City as requiring "effective integration of physical, digital and human systems in the built environment to deliver a sustainable, prosperous and inclusive future for its citizens" [2]. An alternative broad definition, provided by the UK Department for Business, Innovation and Skills (BIS) in which the concept of a Smart City is defined as a dynamic process that makes cities more functional, resilient and adept when reacting to new challenges [3].

Data-driven definitions: IBM defines a Smart City as "one that makes optimal use of all the interconnected information available today to better understand and control its operations and optimise the use of limited resource" [4]. The definition of a smart city provided by Cisco can be also used as another example in which data-

driven contributes as the main factor. According to this definition, “smart cities are those who adopt scalable solutions that take advantage of information and communications technology (ICT) to increase efficiency, reduce costs, and enhance quality of life” [5].

Citizen-focused definitions: According to Smart Cities technical report [6], UK citizens tend to consider smart cities as those who provide them with a decent quality of life, clean environment, and good transport connections. They also associate some other terms with smart cities such as: “technology”, “connected”, “internet” and “modern”.

Smart Cities are presumed to support a foundation for Information Communication Technologies (ICT), data capture and knowledge generation towards economic regeneration, greater social cohesion, infrastructure administration and improvement of governance. According to [7], the top five Smart Cities in 2016 in terms of their smart initiatives in technology adoption, transport services, energy efficiency, Open Data provision and potential economical impact are Singapore, Barcelona, London, San Francisco and Oslo.

It is easy to find accounts of grand aspirations and new initiatives, but rare to hear of genuine achievements beyond relatively small scale pilot projects focusing on specific services. We believe in the Smart Cities dream, but recognise that it will take considerably more than fine words and a sprinkling of flavour-of-the-month technology to make it a reality. It is important that those involved in Smart Cities-related projects cooperate and share experiences and resources to avoid wasteful duplication of effort and repetition of the same mistakes, and to maximise progress towards the common goal.

Despite the benefits of the numerous Smart Cities initiatives, it is difficult to obtain an overview of Smart Cities projects and activities. To begin to address this issue, we have been conducting a survey of Smart Cities projects and data resources focusing six major UK cities (London, Birmingham, Manchester, Glasgow, Bristol and Milton Keynes). The six UK cities selected for this research were identified based on three criteria given from a similar study [8] including: i) Having a large number of funded Smart City projects and programmes, ii) Promoting Open Data initiatives, iii) Providing access to Open Datasets for the public domain.

To provide a basis for analysis and comparison, we have looked into the Open Data initiatives associated with the impact domains across aforementioned cities. The UK’s Department of Business Innovation and Skills report (BIS) [3] has grouped smart solutions into five different categories: Water, Energy, Transport, Waste and Assisted Living Solutions. McKinsey report in 2013¹ identifies seven sectors of the

greatest revenue potentials for urbanisation that are associated with wide-ranging challenges for cities creating complex pressures on City Environments, Infrastructure, Buildings, Networks, Resources and People. Similarly in Smart Cities technical report [6], the proliferation of Smart City programmes and projects are designed to addressing a broad range of city challenges, and bringing opportunities for smarter sectors such as, Governance, Economic Development, Citizenship Living, Environments, and Mobility.

After careful review of aforementioned urban datasets, we have produced ten impact domains (general themes) with associated sub-domains for grouping Smart Cities Open Data and their initiatives, which are aligned with existing literature discussed above.

1. Art & Culture

- Entertainment & Young People
- Travelling & Tourism
- Sports
- Cultural identity & Recreational Programs
- Biodiversity

2. Built Environment

- Infrastructure
- Mapping
- Planning
- Towns & Cities
- Land Use
- Buildings
- Green living space
- Networks and Sensors
- Locations

3. Business & Economy

- Finance
- Economy

4. Community

- Society
- Employment & Unemployment
- Crime & Community Safety
- Housing & households
- Neighbourhood amenity

5. Energy

- Energy Consumption & Carbon
- Wind Energy
- Solar Energy

6. Education

- Education & Schools

7. Governance

[mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information](http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information)

¹<http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information>

- Transparency
- Government Spending
- Elections
- Census Data
- Demographics

8. Health & Well-being

- Child care
- Social care
- Health
- Hospital Admission & Medical data

9. Nature Environment

- Water Quality
- Weather
- Air Quality
- Waste & Recycling
- Resources

10. Transport & Mobility

- Traffic
- Parking
- Public Transport
- Pedestrian

A key aspect of Smart Cities is the collection and analysis of large bodies of data in order to improve the lives of citizens, deliver services more efficiently, and stimulate and support business innovation. Accordingly, we identified as many static data resources as possible that have been established by, or are associated with, the six cities, and classified them according to these domains. Similarly, we analysed the cities' major projects and initiatives to understand to what extent they address the same domains. In this paper, we will use the adjective "static" to distinguish stored data resources from "dynamic" data, i.e. live readings that apply only to the current point in time. We are particularly interested in cumulative stores of data points, a body of data, typically covering an extended period of time, that is available for analysis to reveal trends and other information and insights. This is in contrast to dynamic data that applies only to the current point in time.

Our purposes in publishing the results of the survey, are to:

- make researchers and application developers aware of the existing static data resources they can analyse to gain new insights, or use to prototype and evaluate innovative service ideas;
- promote co-operation among projects with compatible objectives in different cities;
- allow an assessment of collective progress towards realisation of the Smart City vision;
- facilitate roadmapping of research and innovation initiatives so as to incrementally close the gap

between the current state of the art and the ultimate vision;

- encourage similar surveys in other countries using the same ten impact domains.

To complement the paper, we have established a data catalogue² with links to the surveyed data resources and projects, starting with Birmingham, that can be maintained to provide an on-going community resource.

The structure of the paper is as follows. We first review Open Data and its characteristics in Section 2, which is followed by a background of the technical topics and points of reference that are relevant to the Smart Cities data platforms in Section 3. We then analyse the static Open Data sources on a city-by-city basis in Section 4, which is followed by Section 5 that discusses Open Data initiatives associated with the impact domains across aforementioned cities and presents new findings and identifies the gaps and challenges. We finally conclude the work and outline the future direction in Section 6.

2. OPEN DATA

A smart city is not only about technology but also about data and more specifically open data. Accordingly, in a parallel argument to Smart Cities, the concept of Open Data has been emerged. According to Open Knowledge Foundation [9], "Open data is data and content that can be freely used, modified and shared by anyone for any purpose". Open Data Institute³ also defines Open Data as "data that is made available by organisations, businesses and individuals for anyone to access, use and share". Open Data and the development of Smart Cities are paving new openings both at the social and economic levels. Smart Cities can benefit from open data for innovations in:

- Improving government, transparency, and services;
- Solving public problems, data-driven engagement and assessment;
- Empowering citizens through decision making and co-creation;

According to Ojo et al. [8], Open Data initiative, are part of the efforts by governments at all levels to enhance transparency, better empower citizens, foster innovation and reform public services [10]. For example, in 2014, the direct economic benefits from Open Data or public sector information in the European Union was valued around 40 billion and over 140 billion annually across the whole EU26 economy [11].

Since Open Data plays a crucial role to the development of smart cities, we summarise main characteristics of Open Data given from [12] as follows.

- *Completeness or accuracy* of the data made available. Most of the datasets released represents

²<http://birminghamdata.org>

³<https://theodi.org/>

a sample of the data which the user can work with the available attributes and attribute values. Ideally public data should be released in its entirety to the public as this data does not have privacy or security issues. The sense of completeness can only be attained when entire electronic copies of conformant, consistent, credible, processable, relevant and timely public data are provided to the user [13].

- *Availability (accessibility)* cities' open Data should be available in non-proprietary, open formats either through public data websites or other means. The data should be accessible by automated tools with no restrictions imposed on data access. Open data availability and distribution can play an important role in strengthening openness, completeness and accountability thus is considered an important element of Smart Cities initiatives.
- *Usability* it is vital for open data when published to be usable to attain carefully studied objectives with effectiveness, competence and satisfaction in a quantified context of ingestion. Minimal processing should be required before the data is usable. Open data should be provided in formats that promote analysis and reuse, therefore datasets should be cleaned and customised before being published [14].
- *Non-Proprietary* in the sense that data has no legal implications surrounding its use with no entity having exclusive control or proprietary rights over the data with access allowed without registration. Also, proprietary formats require for data to be ordered and stored according to a particular encoding-scheme thus adding unnecessary restrictions over who can use the data, how it can be used and shared, and whether the data will be usable in the future. Accordingly, its may be necessary to make the data available in multiple formats or put the data in a format that allows for it to be easily translated from one to another. Therefore, non-proprietary is a characteristic that Open Data are expected to have.
- *Non-Discriminatory* involves that released open data should not be prejudiced towards a group, community, belief or area. The availability of data should not be restricted to a city or country, a particular religion, race or community. Open data should be available for access without any restrictions or prior registration. Anonymous non-discriminative access to the data must be allowed, including access through anonymous proxies.
- *Persistence* should be made available at a stable Internet location indefinitely and in a stable data format for as long as possible.
- *Documentation* refers to column definitions, names and values including descriptions of how data where collected and cleansed. Open data format and meaning documentation goes a long way to

making the data useful.

3. BACKGROUND OF RELEVANT CONCEPTS

This section discusses the relevant concepts to Open Data in Smart Cities.

3.1. Accessibility and Usability

With the launch of Open Data government scheme, data portals are now providing a single point of access to a growing range of data from institutions and other bodies published by government departments and agencies, public bodies and local authorities. This data is mainly freely available to use through various licenses.

The effective publication of Open Data implicates the application of specialist skills and knowledge in areas such as Open Data licensing and machine processable data and meta-data presentations. Open Data Institute (ODI)⁴ acknowledges the importance to assess and recognise sustainable publication of quality open data. Accordingly, they have developed and maintained a free online tool for open data certification. Certification levels are shown in Table 1.

3.1.1. Open Data Licenses and Certificates

The majority of published datasets are openly available – referred as “Open Data” – for access and used by anyone. Although these datasets are publicly available, they are published under particular licenses, which may restrict their usage for various applications. Some of the main licenses are: UK Open Government, Creative Commons Attribution, Non-Open Government License, Local Auth-With Rights, Open Government License (OGL).

3.1.2. Semi-close, Close, and Confidential Data

Another type of the urban-related datasets are “Controlled Data” that have access limited to approved uses by the publishers to minimise any risks of disclosure.

One of the main publishers is electronic Data Research and Innovation Service (eDRIS)⁵. Their data are held within a highly secure computing environment. eDRIS may agree to permit others to have access for specific research purposes where there is a public interest case for them to do so. This service is only available to UK-based users for UK data sources. From a city perspective, the main data providers are usual governments, local authorities and sector specific enthusiasts.

3.2. Data Quality

Open data is most valuable when it is current, accurate and accessible. There is a universally recognised Five

⁴<https://certificates.theodi.org/en/>

⁵<http://www.isdscotland.org/Products-and-Services/EDRIS/>

	Openly Licenced	Available With No Restrictions	Accessible and Legally Reusable	Machine-Readable	Ongoing Support From Publisher	Regular Updates	Documentation	Uses Unique Identifiers	Communications Team Offering Support
Bronze	X	X	X						
Silver	X	X	X	X	X				
Gold	X	X	X	X	X	X	X		
Platinum	X	X	X	X	X	X	X	X	X

TABLE 1. ODI Open Data Certification Levels

Stars (5★) schema proposed by Sir Tim Berners-Lee for the deployment of Open Data⁶.

- ★ : available on the web (whatever format) but with an open license.
- ★★ : available as machine-readable structured data (e.g. Excel instead of image scan of a table)
- ★★★ : as (2) plus non-proprietary format (e.g. CSV rather than Excel)
- ★★★★: all of the above plus, use open standards from WC3 (e.g. RDF (Resource Description Framework) and SPARQL (SPARQL Protocol and RDF Query Language) to identify things, so that people can point at your stuff
- ★★★★★: all of the above, plus link your data to other peoples data to provide content.

3.3. Data Format

Data are presented in various formats by different organisations within cities. Majority of presentations are both human and machine readable. Some of them can only be analysed by computing programmers, and the rest are easily read by human but unstructured for machines. The major data format presented by different UK cities are:

- CSV (Comma Separated Values)
- XLS (Microsoft Excel Spreadsheet)
- XLSX (Microsoft Excel Open XML)
- XML (Extensible Mark-up Language)
- JSON (Java Script Object Notation)
- PDF (Portable Document Format)
- HTML (Hypertext Markup Language)
- WMS (Web Map Services)

Apart from major formats that are listed above, a few number of urban datasets of the aforementioned UK cities, are also reported in other uncommon formats, including RDF, JPG, WFS, DOC, ZIP, ASPX, SHP, KML, RTF, TXT, WCS, and Atom Feed.

3.4. Data Structure

Open data can be generally categorised into static (i.e. stored) data and real-time (dynamic) data. The main characteristics of each are explained below.

3.4.1. Static Data

Static Data can be further divided into structural or contextual data that describes aspects of the world not changing so much over time, such as train station and bus stop locations, and cumulative or historical data. The latter has a temporal dimension, consisting of data points recorded over an extended period, allowing trends and correlations to be studied. In open data portals, contextual datasets are typically downloaded in a one-time operation; however, they might still be periodically updated (e.g. yearly or a few times a year). Historical data is likely to be published as a series of datasets covering a sequence of time periods (e.g. month-by-month). This may be referred to as “incremental data”.

3.4.2. Dynamic Data

In contrast to structural or contextual data, dynamic data, i.e. live readings that apply only to the current point in time and also referred as real-time, is constantly changing like traffic congestions and number of available parking spaces. Smart Cities projects often involve constructing an Internet of Things (IoT) by instrumenting the city with networks of smart devices producing streams of data describing the current state of the city and its infrastructure. This can be made available in real time for access on-demand via application programming interfaces (APIs), and also stored in cumulative datasets for analysis.

3.5. Metadata

Open data metadata is a controlled vocabulary that describes datasets and APIs, making it easy to discover, locate, retrieve use or manage information from datasets. The challenge is to define and name standard metadata fields so that data users have sufficient information to annotate, understand the semantics and process the described data. The more knowledge that can be conveyed in a formal representation, the more valuable data becomes.

Establishing an agreed vocabulary is the key for communication. Table 2 [15] depicts the common core metadata models used by main Open Data portals. These metadata models are based on Dublin Core⁷ as a basic layer for metadata standards. It is the

⁶<https://www.w3.org/DesignIssues/LinkedData.html>

⁷<http://dublincore.org/>

Open Data Management Platform	Metadata Model
CKAN	CKAN
Open Data Soft	DCAT
SOCRATA	Socrates
DKAN	DCAT, INSPIRE
ArcGis Open Data	INSPIRE
Esri Geoportel Server	Open Geospatial
Junar	DCAT, INSPIRE

TABLE 2. Metadata Models for Open Data

specific layer of a metadata standard that determines the subject-specific use of the standard, while the distribution layer is the one which makes the metadata model for each Open Data portal.

The most commonly used metadata standards is (DCAT) Data Catalogue Vocabulary. DCAT is a hierarchical vocabulary defined in three levels of metadata elements: *Required*, *Required-if* (conditionally required), and *Expanded* fields. These elements were selected to represent information that is most often looked for on the web. To assist users of other metadata standards, mappings to equivalent elements in other standards (e.g., Dublin Core) are provided. Geographic metadata standards *ISO 19115* is an internationally-adopted schema for describing geographic information and services. It provides information about the identification, extent, quality, spatial and temporal schema, spatial reference, and distribution of digital geographic data.

3.6. Open Data Management Platform

Each smart city has a data management platform, that is used for publishing and exploring urban data. Some of the major platforms as follows.

*CKAN*⁸

The Comprehensive Knowledge Archive Network (CKAN) is a web-based open source data catalogue system that is mainly used by public institutions seeking to share their data with the general public. It has a modular architecture that allows for extensions to be developed to provide additional features such as harvesting or data upload. CKAN uses its internal model to store metadata about the different records, and presents it on a web interface allowing users to browse and search this metadata. It also offers a powerful API that allows third-party applications and services to be built around it.

CKAN is used in many UK cities, including: Birmingham, Glasgow, London and Manchester.

*DKAN*⁹

DKAN is a Drupal-based Open Data tool for cataloguing, publishing and visualising data to the public. DKAN is maintained by GovDelivery and was

⁸<https://www.ckan.org>

⁹<http://getdkan.com>

designed for groups who want to share data with the world in open, public, machine-readable formats, or generate maps, charts and other visualizations from the data. The DKAN Dataset and Datastore can also be added to existing Drupal sites to add data portal functionalities.

*Fiware*¹⁰

FIWARE is an open cloud-based European middleware platform, driven by the European Union, with the aim to develop Smart Applications for the Future Internet. The API specification of FIWARE is open and license-free with the objective to facilitate cost-effective development and distribution of Future Internet applications and services in diverse application domains, including Smart Cities, sustainable transport, logistics, renewable energy, and environmental sustainability. Milton Keynes uses the new FIWARE business framework for the TM Forum APIs¹¹.

*Socrata*¹²

Socrata is a cloud-based platform that enables public sector organisations to easily manage and share their data. This platform makes the data accessible to anyone with a web connection, while also giving them the power to visualize and analyse the data. Socrata Cloud-based Data Democratization solutions for Government (CDDG) are designed and purposely-built to meet strict government standards. It provides a software-as-a-service (SaaS) data platform and cloud applications exclusively for city, county, state and federal government organisations to be discoverable, usable, and actionable for government knowledge workers and the citizens they serve.

*OpenDataSoft*¹³

OpenDataSoft allows restricted and open sharing ecosystems for cities, public organisations and companies data. It specializes in transforming structured data into API and visualizations. Its interface is intuitively designed to empower anyone, regardless of technical skill level, to upload easy-to-understand Open Data, or to even share data within an administration or organisation of any size. Bristol uses OpenDataSoft for smart technologies and digital connectivity to meet the city's environmental, social and economic Open Data challenges for Smart Cities.

*Swirrl*¹⁴

Swirrl is a small technology company that provides Open Data publishing solutions to make data easy to find, understand and reuse. Swirrls PublishMyData

¹⁰<https://www.fiware.org>

¹¹<https://www.fiware.org/>

¹²<https://socrata.com>

¹³<https://www.opendatasoft.com>

¹⁴<http://www.swirrl.com/>

platform is used by public and private sector groups to publish and manage their data online in structured, machine-readable forms. Swirrl helps publishers to meet open standards from the World Wide Web Consortium (W3C). It creates custom browsing and discovery features, such as visualisations and applications, to ensure complex data can be used by non-technical users.

*Hypercat*¹⁵

Hypercat was developed as a solution to the IoT interoperability and data sharing needs of the Technology Strategy Board IoT Demonstrator Programme. It is designed for uncovering information about IoT resources over the web. It allows for servers to provide sets of resources to clients, enriched with semantic annotations to allow for interoperability. Users are free to choose or formulate sets of annotations to suit their specific needs. With Hypercat developers can write applications that will work across many servers, thus breaking down the walls between vertical silos. Where implementers choose similar or overlapping semantics, the possibilities for information sharing and data integration are increased.

4. DATASETS & SERVICES OVERVIEW

This section discusses the urban datasets submitted by six UK cities: London, Birmingham, Manchester, Glasgow, Bristol and Milton Keynes. The data are collected during the period of **February to March 2017** and are based on the secondary information from data publishers such as City Council websites, innovative project representative publishers and other major data hubs, which include a range of urban-related data, such as commercial, governmental, transportation, citizen engagements and etc.

For the initial analysis of the such datasets, we have first looked into major data publishers and themes for each city. This overview is then followed by the secondary analysis in Section 5 that we have grouped these datasets into 10 major themes and discussed the proportions.

A. London

London is the capital of England and the most populated city of the United Kingdom, as well as the most populous municipality in the European Union. This metropolitan city has been divided into three major areas: Greater London with the population of 8,673,713, Urban region with the population of 9,787,426, and Metro area populated by 13,879,757 people in 2015 according to the report [16].

For this study, we have taken into account three main sources: London Datastore¹⁶, UBDC¹⁷,

DATA.GOV.UK. An overview of the datasets and data publishers for London are as follows.

Data Publishers (total number: 22)

Greater London Authority (GLA)
Office for National Statistics (ONS)
Transport for London (TfL)
Census Information Scheme
Department for Communities and Local Government (DCLG)
Department for Education
Department of Health
London Fire and Emergency Planning Authority
Department for Work and Pensions
Department for Transport
HM Revenue & Customs
Department of Energy and Climate Change
Valuation Office Agency
Department for Environment, Food and Rural Affairs
London Borough of Camden
London Borough of Barnet
Marine Environmental Data & Information Network
Environment Agency
London Borough of Hackney, Hounslow
City of London
United Kingdom Hydrographic Office
Scottish Government Spatial Data Infrastructure

Main Themes

- **Demographics:** This sector presents datasets about the reports, population and census data on various topics for various regions by different periods (2001-2015), such as household projections, ethnic groups and religion by Ward, age and gender, children in poverty, facilities in London, census and key statistics etc. These also include a summary of demographic and related data for each Middle Super Output Area in Greater London. Datasets are mostly presented in CSV, XLS and HTML.
- **Employment and Skills:** This category consists of reports on different topics such as employment and unemployment's rates per Borough by gender, age, status, ethnicity, disability, full-time or part-time jobs, job seekers allowance claimants etc. Datasets are mostly presented in CSV and XLS.
- **Environment:** This category consists of datasets with the focus on various environmental topics, such as domestic and industry waste and recycling scheme, different types of energy consumption, air quality and pollution, Environmental Footprint Data, Climate Change Mitigation and Energy Annual (2009-2012), Low Carbon London Electric Vehicle Load Profile, London's Urban Heat Island etc. Datasets are mostly presented in CSV and XLS.

¹⁵<http://www.hypercat.io/>

¹⁶<https://data.london.gov.uk/>

¹⁷<http://ubdc.ac.uk/>

- **Transparency:** This sector contains the local government transparency reports, such as, Fire Brigade, Mayor's Crowdfunding Programme, annual London survey, various local and council election results, GLA key performance indicators, investment programs for different periods etc. Datasets are mostly presented in PDF, CSV and XLS.
- **Housing:** This category contains datasets related to housing matters and their residences in London and its regions. Some of them are: GLA affordable housing programme, local authority average rents, registered social landlord housing stock, vacant dwellings, homelessness provision, DCLG affordable housing supply, average private rents, town centre locations, council tax charges and bands, The economics of London's housing market and housing sales, overcrowded households and empty homes by Borough. Datasets are mostly presented in PDF and XLS.
- **Health:** This category includes the reports and information about health matters in London and Borough. Some of the topics are: Smoking indicators, Sports participation rates, Immunisation rates for children, Births and fertility rates, Health inequalities strategy indicators, Hospital admission rates, Legal abortion rates, Breast cancer screening, Country of birth census, Life expectancy by Ward, Number of registered blind and partially sighted people, Hospital admissions due to injury of children, Immunisations against influenza, Subjective personal well-being, Road casualties by severity, Obesity in adults, Census health & care, Drug treatment rates, Life expectancy by deprivation, Quality and safety hospital audits, Number of people Registered as deaf or hard of hearing etc. These datasets are mainly in XLS format.
- **Transport:** This category contains the datasets related to transportation. There are a significant Open Data available for London and its regions including: London Underground Performance Reports, Public Transport Accessibility Levels, Number of Bicycle Hires, Busiest Airports by Passenger Traffic, Train Station Usage, Walking and Cycling, Traffic Flows, Type and Number of Licensed Vehicles, Road Transport Energy Consumption, Flights and Passengers from London Airports, Bus Stop Locations and Routes, Census Key Statistics for Cars and Vans, TfL Live Traffic Disruptions, TfL Live Roadside Message Signs, River Boat Timetables, Oyster card journey information. The datasets are mostly in XLS and CSV formats.
- **Business and Economy:** This category contains datasets with the focus on local and regional businesses for London and its areas including: Jobs and Job Density, Number of International Visitors to London, Business Demographics and Survival Rates, Local Units and Enterprises by Employment Size, London Business Survey and Profile, Commercial and Industrial Floorspace, Tourism Spend Estimates, London's Food Exports etc. The reports are presented in CSV, XLS and PDF.
- **Education:** This category contains datasets with the focus on education and children living in London and Borough. Some examples are: Students by higher education provider, Pupils with Statements of Special Educational Needs, Schools and Pupils by Type of School, Participation of Under-Represented Groups in Higher Education, School Revenue, London Annual Education Report, Results by Free School Meal Eligibility, Characteristics of Pupils Living, Physically Active Children.
- **Other important categories:** Greater London authority has a large collections of Open Data in other categories such as Planning, Crime and Community Safety, Young People, Sport, Young People, Arts and Culture, Income, Poverty and Welfare.

B. Birmingham

Birmingham is the second largest city after London in the UK with a population of 1,101,360 in 2014 according to the census and population report by Birmingham city council¹⁸.

Open datasets related to cities data are collected from three main resources including: DATA.GOV.UK, Birmingham Data Factory¹⁹, and West Midland Open Data²⁰. Among these, the dominant data formats are CSV, XLS, WMS and HTML.

Data Publishers (total number: 9)

Birmingham City Council
 Centre for Ecology & Hydrology
 Centre for Environmental Data Analysis
 South Birmingham Primary Care Trust
 Office for National Statistics
 NHS Sandwell and West Birmingham
 NHS Birmingham South and Central
 NHS Birmingham Cross City CCG
 Natural England

Main Themes

- **Travel and Transport:** It is the main category that includes data related to Birmingham real-time traffic data, cycling routes and sensors, bus, rail and metro user profiles and midland passengers counts. The datasets are mostly represented in CSV format. An API is provided for accessing Urban Traffic Management and Control database.
- **Council Business** It is the second top category that consists data related to purchase cards transactions, taxi, rail and accommodation costs, payment information statistics, individual incidents

¹⁸<https://www.birmingham.gov.uk/census>

¹⁹<https://data.birmingham.gov.uk>

²⁰<http://opendatawestmids.org.uk>

(2014), and council tax arrears by stage (1997-2012). The data are mainly presented in CSV and XLS formats.

- **Environment:** The next category is environmental data, which consists of data related to waste, street cleaning schedule, energy consumption from Birmingham Data Factory. In addition, UK Government Data also reports on a number of environmental datasets on Birmingham tree preservation orders, riparian habitats and wildlife corridors, Continuous measurements of rainfall and rivers (2007) etc.
- **Your Local Area and Leisure:** This category includes datasets in relation to housing, urban living environment and Leisure activities, such as Social Housing Aggregated, Population and Census Data, ID Homelessness indicator (2007), Birmingham Civic Dashboard, Green Spaces, Parking zones, Domestic Waste Centres, Statutory Common Land.
- **Towns and Cities:** This category consists of information about cities and towns around Birmingham. Some examples are: High Density Temperature and Meteorological measurements within the Urban Birmingham Conurbation.
- **Education:** Most datasets in this category include information about the education matters such as, Birmingham Schools, Secondary School Admissions intake in 2013, 2014.

C. Manchester

Manchester is a major metropolitan city with a population of 530,300 in 2015 according to Office for National Statistics ²¹. A recent EU report [17] ranked Manchester as the fifth most successful Smart City amongst 240 EU28 cities.

The datasets for Manchester are collected through three sources: Manchester city council via UBDC, DATA.GOV.UK, and Greater Manchester Public datasets (DataGM²²). UBDC lacks the organisational representation of the datasets with appropriate links to their categories and publishers as we could not feature their categories. Majority of datasets belong to the two later resources, that are not properly classified, rather than being tagged with particular keywords.

Data Publishers (total number: 10)

Environment Agency
 Transport for Greater Manchester
 Manchester City Council
 Salford City Council
 Centre for Environmental Data Analysis
 Crown Prosecution Service
 Centre for Ecology & Hydrology
 British Geological Survey

United Kingdom Hydrographic Office
 Trafford Council

Main Themes

- **Environment:** Almost half of the datasets from DATA.GOV.UK are in this category, such as Water Quality Archives for greater Manchester in different years (2008-2015), Sites of Biological Importance in greater Manchester, Air quality monitoring, Soil and Ecosystem functions, Tree Preservation order. However, not all the datasets are related to Manchester city but other UK regions. The dominant formats are XML and HTML files with links to WMS.
- **Transport:** These datasets are related to Great Manchester cycle, railway, bus, car, traffic signals, routes, stops and parking. There are also datasets related to transport statistics from other cities in UK and road casualties in Manchester. These reports are in XLS, CSV and PDF.
- **Towns and Cities:** Like Bristol and Birmingham, this category contains the datasets related to the residences, their lifestyles and activities. Various topics are covered such as, Licensed Premises, Social Housing Asset, Manchester Airport and Mineral Safeguarding Zone, Cemeteries, Compulsory Purchase Orders, Town and Neighbourhood Centres, and Green spaces, presented in HTML and WMS.
- **Court & Justice:** To compare with other cities, this category is unique for Manchester that contains monthly publication of criminal case outcomes in the magistrates' courts and in the Crown Court by principal offence category and by CPS Area (2010-2013), and presented in CSV format.
- **Government:** This category is also specific for this city, which consists of Manchester governmental matters such as local election results, administrative service, senior salaries, pay policy statement, constitution, and TfGM Fraud Investigation reports in CSV.
- **Mapping:** Similar to other cities, this category includes the information about local lands, properties and assets data shown in HTML and WCS format.

D. Glasgow

Glasgow is the largest city in Scotland, and the third major city in the UK with the population of 606,340 in 2015 ²³. The open datasets collected for this city are from three main resources: Data Glasgow²⁴, UBDC and DATA.GOV.UK. The limitation with Data Glasgow datasets is that there are many overlaps between the tags such as, health and Health, Hospital and hospital admission, transport and Public transport. However,

²¹<https://www.ons.gov.uk/>

²²<http://www.datagm.org.uk/>

²³<http://www.understandingglasgow.com/>

²⁴<https://data.glasgow.gov.uk/>

the datasets are rich and contain various categories.

Data Publishers (total number: 11)

Glasgow City Council
 Urban Big Data Centre
 Scottish Government Spatial Data Infrastructure
 Centre for Ecology & Hydrology
 British Geological Survey
 United Kingdom Hydrographic Office
 Marine Environmental & Information Network
 Joint Nature Conservation Committee
 Department for International Development
 Scottish Government
 Food Standards Agency

Main Themes

- **Health:** It is the major category that includes Scottish health survey, fire and rescue service, police divisions, various source of noise contours such as airports, railways, roads, industry and NHS health boards. The datasets format are varies from XML and HTML to WMS.
- **Environment:** This is the second category that consists of live data page for energy and water consumption, habitats and species points record (1970), urban soil chemistry data, mineral resource polygons central belt of Scotland etc. Most datasets are in HTML or CSV formats accompanied by WMS presentation.
- **Towns and Cities:** Datasets in this category are related to Glasgow council and city matters, which cover urban areas locations and settlement in Scotland (2010, 2012), heat demand in different years and landscape character presented in both HTML and XML.
- **Transport:** This category includes information about Glasgow bridges, winding holes, boat lift and truck road network for Scotland shown in HTML, XML, and WMS.
- **Mappings:** Glasgow city council has also collected the data related to location of land registration points, marine management, and broad rental market areas. These datasets are mainly shown in WMS format.

E. Bristol

Bristol is located in South West England with an estimated population of 449,300 in 2016²⁵. It is England's sixth and the most populous city in Southern England after London.

Open datasets for Bristol are collected from two main sources: Open Data for Bristol²⁶ by the city council and the DATA.GOV.UK hub. Bristol city council uses OpenDataSoft platform to present their datasets, which contains other types of data formats such as, Data

Lens pages, Stories, Datasets, Maps, Calenders, and Forms. A list of data publishers and main themes from DATA.GOV.UK as follows (Figures 1, 2).

Data Publishers (total number: 10)

United Kingdom Hydrographic Office
 Marine Environmental Data & Information Network
 Marine Management Organisation
 Bristol City Council
 Centre for Environment, Fisheries & Aquaculture Science
 Centre for Ecology & Hydrology
 Joint Nature Conservation Committee
 Centre for Environmental Data Analysis
 Natural England
 Environment Agency

Main Themes

- **Towns and Cities:** These datasets mainly contain the number and location of public and residential facilities, green space, allotment plots, polling stations, recycling , waste and litter bins, wireless hots post and parking schemes. The data are mostly presented in HTML, WMS and WFS formats.
- **Government:** This category contains data about important issues facing the council and city, such as, elections, events, committee meetings etc. These datasets are presented in both CSV and RDF.
- **Environment:** Majority of datasets generally consist of information about Bristol air and water quality, river and rain measurements, water body status, energy consumption and marine renewable energy, habitats points records and wildlife corridor , shown in CSV format.
- **Transport:** Datasets include information about Bristol's bus stops, car parks, cycling routes and pumps, traffic survey and regulation reported on HTML, WMS and CVS.
- **Health:** Bristol city council has published 26 datasets in the context of health. The topics are mainly about NHS Bristol reports of various health matters, census-health of different years, life expense and mortality, and healthy lifestyles. The information are presented in table format.
- **Finance:** There are table datasets published by Bristol city council that report on occupations and census-economic activities in Bristol (2011), income support claimants (2012) and payments over 500 (2011, 2012, 2013).

F. Milton Keynes

Milton Keynes is recognised as the best example of Smart City transformation across the UK. Its fast growing economy set to grow 64 percent by 2026²⁷.

²⁵<https://www.bristol.gov.uk/statistics-census-information/>

²⁶<https://opendata.bristol.gov.uk/pages/home/>

²⁷<https://www.tmfforum.org/resources/research-and-analysis/smart-cities-enabling-economy-data/>

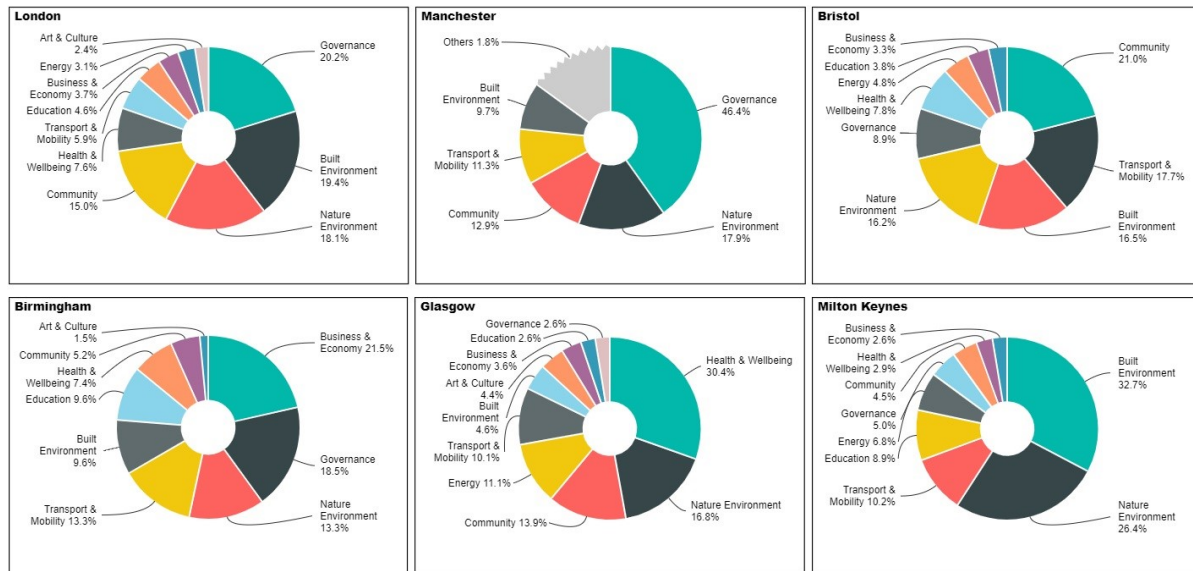


FIGURE 1. Open Static Dataset Distribution Across Major Themes By Six UK Cities

MK:Smart data hub²⁸ is specially designed to provide datasets related to this city, in total 410 datasets, that are presented in XLS, XLST, CSV, XML and JSON formats.

Data Publishers (total number: 10)

- Milton Keynes Council
- Natural England
- Canal & River Trust
- Vale of White Horse District Council
- South Oxfordshire District Council
- Joint Nature Conservation Committee
- NHS Milton Keynes CCG
- Milton Keynes Primary Care Trust
- Milton Keynes Hospital NHS Foundation Trust
- Marine Environmental Data & Information Network

Main Themes

- **Transport:** MK data hub has published Milton Keynes transportation information in this category including: Bus location and passenger count, Traffic busyness, Live traffic information, Roundabouts From Open Street Map, Parking and Charging Bays.
- **Sensors:** This is a new category to compare with other cities. Milton Keynes collects data from various sensors around the city region. Some of the datasets are Soil moisture sensor from Koubachi device, Frequency of tickets issued from parking meter machines, netatmo weather station from temperature, humidity, pressure, water level and public recycling for mixed and glass items.

- **Environment:** Similar to other cities, this category contains datasets related to Milton Keynes environmental matters such as, Collection of satellite and aerial imagery, Weather forecast, OpenWeatherMap from current weather, Planning Applications, Parks From Open Street Map.
- **Health:** This category contains datasets with health and well-being information. Some examples are: Independent Food Retailers, Doctors Surgeries From Open Street Map, Hospitals From Open Street Map and Disability Living Allowance, Census 2011 Health.
- **Business:** This category, similar to other cities, includes datasets with business data and planning in Milton Keynes such as, OpenCorporates companies registered, Post Offices From Open Street Map, Milton Keynes Sports Centres From Open Street Map, Food Establishments Info and Ratings, Census 2011 Economic activity of people in Milton Keynes wards, Census 2011 Occupation by Milton Keynes wards.
- **Education:** This category contains datasets with school and education information collected in 2011: School deprivation, workforce school and college data, spend per pupil school and college data, pupil absence school, Playgrounds From Open Street Map, Libraries From Open Street Map, Schools From Open Street Map, Skills for Life Survey Numeracy, Literacy - Qualifications in Milton Keynes wards, Milton Keynes Museums Information From Open Street Map. These datasets are presented in CSV.
- **Crime:** MK Data Hub also has a collections of dataset with crime issues including Burglaries, Criminal Damage, Robberies, Violence against the

²⁸<https://datahub.mksmart.org/data-catalogue/>

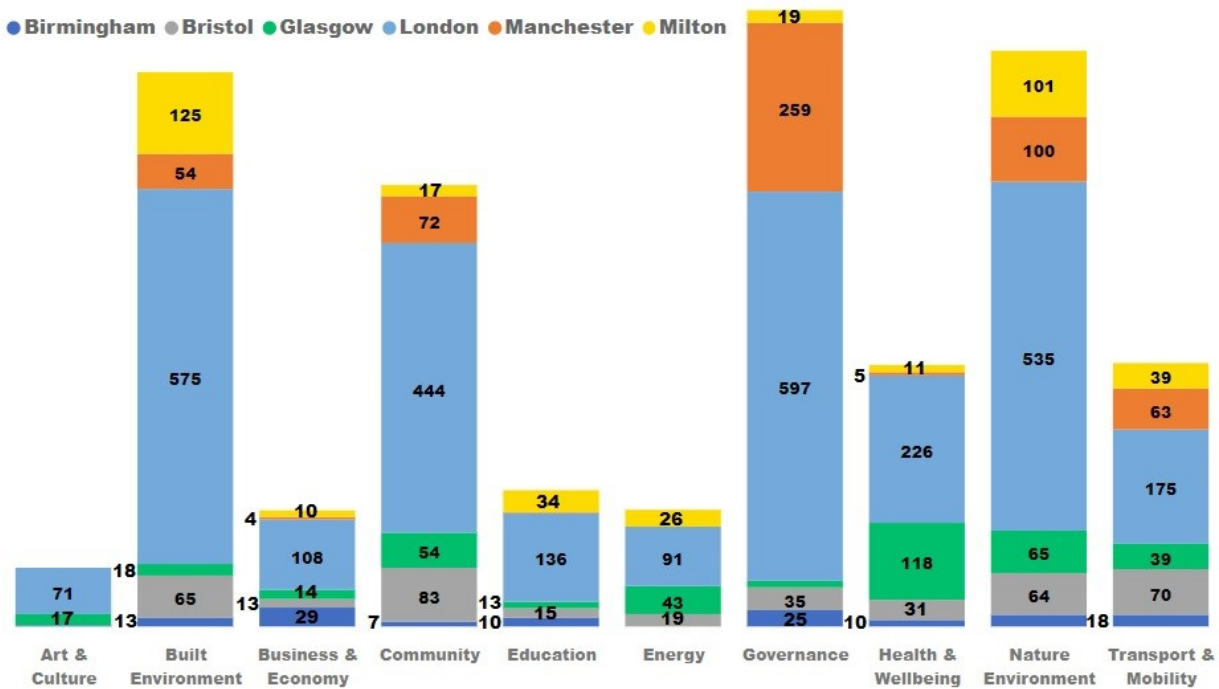


FIGURE 2. Stacked Chart Representation of Urban Open Static Data within Themes - alphabetical order by Six UK Cities

person, crime statistics.

- **Energy:** It is another popular category that MK Data hub provides datasets between 2010-2013 on these topics: Samsung Trial Average Electricity Consumption, Lower Layer Super Output Area (LSOA) gas and electricity consumptions, and fuel poverty data England.
- **Other Categories:** MK Data hub has a few datasets for other categories including: Demography, Employment, Housing, Geo and location, Public and Social Media.

5. FINDINGS & DISCUSSION

Given the initial analysis of the aforementioned cities' Open Data presented in Section 2, we have then used this new impact grouping in the Section 1, to discuss the new findings in this section.

Our further analysis has produced two different representations of these datasets: First, Figure 1 features the datasets' proportions for each city individually. Second, Figure 2 compares the distribution of such datasets within the domain impacts for each city in a *StackedChart* model. These findings are then linked to city visions and relevant actions in the next section.

5.1. City Visions and Relevant Actions

Following our findings from the previous section which are purely based on static Open Data, we have then identified the vision and actions of the aforementioned UK cities towards the impact domains using various

cities' reports and urban project websites. Figure 3 summarises the innovative projects associated with relevant static and real-time datasets, apps and services.

Our findings indicate that among ten major domains, only a few of them are linked to the innovative actions and projects in cities. For example, according to Grayson [18], in Built Environment, Birmingham has become a leading global green city within the West Midlands region. This city has established a Green Commission, adopting an Ecosystem Framework, which results in several strategies and projects to improve city environment as listed in Figure 3. However, the proportion of static Open Data in this domain for Birmingham is only 7% (Figure 1), although we anticipate that these numbers would be much larger for dynamic and real-time data which are accessible by APIs.

Similarly, London, Milton Keynes, Bristol and Manchester are becoming UK-leading cities in energy efficiency. London is engaged in a number of EU funded projects aiming at carbon reduction and decentralised energy [19]. Milton Keynes is also equipped with a range of advanced energy installations, such as the Falcon smart grid, an extensive electric vehicle charging infrastructure, and a district heating system as listed in Figure 3. Bristol has launched Smart City Bristol in 2011, aiming to reduce CO₂ emissions by 45% by 2020 with the use of smart technology. The strategy, led by the council's Smart Cities team, focuses on smart transport, smart energy and smart data [20]. Similarly, Manchester has focused on building

City	Major Smart Projects	Transport & Mobility		Built Environment		Natural Environment		Energy		Health & Wellbeing		Culture & Community		Governance		Used Datasets	Apps, APIs and Services
		Static	Dynamic	Static	Dynamic	Static	Dynamic	Static	Dynamic	Static	Dynamic	Static	Dynamic	Static	Dynamic		
London	Transport for London	X	X	X	X							X	X	X	X	Atmospheric Emissions Inventory Journey Planner timetables Station locations and facilities Tube departure boards, line status & station status WiFi access points Coach parking sites & locations Pier locations Geographic boundary of the GLA road Live Traffic Disruptions Post code areas by major road schemes Live traffic camera images & roadside message signs Licensed private hire operators Cycle Superhighways, Quietway route & hire data Oyster Ticket Stop locations & Card information Step-free Tube guide and toilet data Busiest times on trains and in stations Rolling Origin & Destination Survey London Underground passenger counts data Dial-a-Ride statistics	Unified API London Air API Journey Planner API Beta Cycling data in the unified API Live bus and river bus arrivals API (instant & stream) Public Transport Access Level Tube map service updates Journey planner Roads
	Low Carbon London	X			X	X	X				X	X	X	X	X	Household energy consumption data	Installing Smart Meter Energy Consumption
	Camden Council				X	X				X	X	X	X			Residential dynamic time trial using council data sources	Smart London Innovation Network
	Sensing London						X						X	X		Sensors across five London sites	Smart London Innovation Network
	London Living Lab											X	X			Use cases with communities, city officials & stakeholders	Smart London Innovation Network
	Tower Bridge Idling			X	X	X										The sensors plus information from the bridge management system	Smart London Innovation Network
	Hyde Park Sensing	X	X		X	X	X									Real-time data on the park's eco-system and social fabric	Smart London Innovation Network
	Talk London					X	X				X	X	X	X		Big policy decisions of the future using citizen's opinion	London City Council
Care Connect	X	X					X				X	X	X	X	Enabling people to share your experience, ask a question or report a problem with an NHS service	NHS Choices, myhealthlondon	
Birmingham	Birmingham Smart City Eco-system (Birmingham Smart City Roadmap)	X	X		X		X			X	X	X	X		Using West Midlands Open Data Forum	Birmingham Smart City Alliance (SCA) Birmingham Science City (BSC) Innovation Birmingham Ltd	
	Digital Connectivity in Birmingham (Birmingham Smart City Roadmap)		X	X								X			Working with Birmingham data centre organisations, the London Internet Exchange (LINX) and local developers	Droplet Pay: Smart phone app Go On Birmingham Hello Business	
	West Midlands Urban Traffic Control Major Scheme		X	X		X						X	X		Integrating transport data From Traffic Signal Junctions, Variable Messaging Systems, Environmental Monitoring Systems	Network West Midlands (NWM) NetNav Walkit	
	Smart Spaces		X	X			X	X			X	X	X	X	Including smart metering for significant energy saving in public buildings	Working with social landlords and householders to offer energy efficiency measures in domestic buildings	
Manchester	Eastserve Go On Manchester				X					X	X	X	X		Accessing interactive public services as well as local information and line services	Manchester Digital Development Agency	
	CityVerve BT CityVerve Portal	X	X		X	X	X	X	X	X	X	X	X	X	Platform of Platform technology by creating a secure BT CityVerve Data Hub for acquisition, management and processing of the vast amounts of data	Uniting applications, enabling the intelligent collection, interpretation and use of data through a Restful API	
	Greater Manchester Data Synchronisation Programme (GMDSPP)	X	X		X	X	X	X			X	X	X	X	Allowing the free flow of data between public sector organisations across Greater Manchester	Engaging the digital community and SMEs to encourage the creation of new digital services and insights	
	Transport for greater Manchester (TIGMs)		X					X				X	X		Real-Time Open Data Portal	Providing developers with an API to accessing data from across the region's transport network.	
Glasgow	Glasgow City Energy			X	X	X	X	X			X	X			Using rich data to create a detailed portrait of the city's energy consumption	Energy App BUILDAX SENSOR ELTEC SENSOR	
	Active Travel		X			X				X	X				Based on real user data	Cycling App Walking App Walkanomics	
	Intelligent Street Lighting		X	X	X							X			Adding more control and efficiency to the lighting network while harnessing the power of real time data to improve both lighting and safety throughout the city	Riverside Walkway Gordon Street Merchant City	
Bristol	Smart Spaces (Smart City Bristol)		X	X		X	X	X	X	X	X	X	X	X	Including smart metering for significant energy saving in public buildings	Working with social landlords and householders to offer energy efficiency measures in domestic buildings	
	3e-Houses (Smart City Bristol)		X	X		X	X	X	X	X	X	X	X	X	Monitoring electricity and gas with data sent through a wireless router	Providing participants with latest touch screen Toshiba tablet so they could view their electricity and gas usage online every hour through a visual website	
	City Living Lab		X	X	X	X	X	X	X	X	X	X	X	X	Combining data from different sources	Hosting hack events in order to encourage citizens and businesses to use them	
Milton Keynes	MK.Smart		X	X	X	X	X			X			X		Demonstrating how data gathered from the city's censored infrastructure and other sources	Helping better manage utilities and decrease the city's carbon footprint	
	Smart Street Lighting	X	X	X	X	X	X								Trialling a number of smart initiatives	Installing smart Sensors on streets	

FIGURE 3. UK Cities' Projects, Apps & Services Within Major Themes based on Static & Dynamic Open Data

energy management with a number of strategies and projects, such as building energy management system for Town Hall Extension and Central Library, developing a virtual power plant solution supporting electric vehicles based at Manchester Science Park, creating 3D Visualisations to model real-time energy usage [6].

However, our findings of relevant actions and projects

for each city are shown in Figure 3, is inconsistent with the Open Data distribution in each city shown in Figure 1. The data proportion for Energy sector London, Milton Keynes, Bristol and Manchester are 7%, 7%, 5% and zero percent respectively. Although we have to take into account that these percentages are purely based on static Open Data and their proportions likely would increase by considering the huge volume of

real-time data generating via IoT devices and APIs.

In Transport and Mobility theme, the “Smart London Plan” launched by the GLA in December 2013 is one of the major initiatives within the UK cities. The project has been aimed to use technology to integrate transport data from different agencies such as highways agency, police etc. into a single platform [6]. Different projects (Figure 3) are looking at aggregating disparate data sets and connect sensor networks across London to facilities live data feeds, advanced modelling, simulation, analysis and visualisation [6]. Moreover, London Smart Plan is using the data collected from Oyster cards to understand congestion patterns and plan future investment [19].

By contrast, we see contradicted proportion of Transport (6%) and Mobility for static Open Data in London in Figure 1. Therefore, it is not wrong assumption that the aggregation of different datasets for London transport data are mostly based on dynamic datasets rather than the static data.

On the other hand, Birmingham, Bristol and Milton Keynes have increasing number of static datasets in this sector, 13%, 18%, 10% respectively, which are aligned with their recent innovative projects. For example, in Birmingham, academic and industry partners have used transportation Open Datasets in a number of projects to improve public transport services (Figure 3). Similarly, Bristol authorities have used real-time traffic data collected through the Traffic Control Centre to measure congestion in the city combined with data collected through their Open Data Portal [21], although our findings show that there is no particular action or project planned towards this goal.

In Community theme, Birmingham is moving fast towards improving citizen living. Digital Birmingham is a council-owned partnership organisation that formed a Smart City Commission in July 2012 and launched the Birmingham Smart City Roadmap in March 2014 [22, 23]. According to the city report [6], the roadmap sets out 49 actions grouped under three main themes: Technology and Place (which involves improving broadband connectivity and sharing Open Data), People (focusing on digital inclusion, improving citizens’ ICT skills and implementing new business procurement processes) and economy (mainly around digitising social care, improving energy efficiency and smart mobility) [18, 24]. However, we have found few static Open Data (5%) resources available for this sector in Birmingham (Figure 1). The contradicted figures again confirms the gap between the static Open Data sources and the innovative projects, although it is very possible that these projects are utilise APIs for real-time data.

5.2. Challenges and Solutions

Above, we have outlined some of the inconsistencies between the availability of static Open Data and the

city visions along with their relevant actions for the six UK major cities. These findings cannot be ignored given the increased attention to the data science field with claim that data is the new oil. Data-driven Smart City definitions such as those from IBM and Cisco have often quoted above. However, it is not clear from our study why Smart Cities have not chosen to collect and publish more data following with this new vision. Perhaps the government see more immediate impacts from other initiatives – yet surely this is short-sighted. Some of the challenges are briefly discussed below.

UK cities are facing considerable challenges identified by different cities reports [20, 21, 25], focus on increasing urban populations and its consequences. Such challenges include an increase in infrastructure stress, changing public service needs, and the need for additional employment opportunities. Bristol in particular is expecting a 31% rise in population by 2028 [20]. In contrast, Manchester is already feeling the effects of these changes, having increased in population by nearly 20% in the past decade. Milton Keynes is anticipating an increase in the age of its population, and highlights the need for a focus on the prevention and management of age-related illness in response [21]. In contrast, Birmingham has one of the youngest populations in Europe, and is expected to get younger by 2035 with above-national-average growth in the number of people from all age groups below 65 [22, 23].

Numerous cities around the world are also faced with the challenges posed by population growth particularly with regards to infrastructure upgrades and services provision to enhance their citizens’ quality of life. Some of these challenges are being addressed by developing completely new solutions, such as Vienna (Austria), Mexico Lake (China), Masdar (Abu Dhabi), Lavas (India), and Skolkovo (Russia) [26], which can be modelled by the UK smart cities.

6. CONCLUSION AND FUTURE WORK

In this paper, we have provided a survey that reviews the current status of static Open Data and innovative actions for six UK major cities(London, Birmingham, Manchester, Glasgow, Bristol and Milton Keynes), that are moving towards the Smart City dream. In this research, we distinguish between static datasets, which are accessible via various data hubs, and dynamic data, which are accessible via APIs in real-time. Our reason for this distinction is that we are interested in the potential benefits to cities that can accrue from the analysis of accumulated bodies of data, rather than from the availability of up-to-date data. Therefore, our findings are purely based on static Open Data.

Our aim in this paper was to raise awareness and assess the existing static data resources, to improve the collaboration among projects with compatible objectives in different cities, and also to identify the gaps in coverage that are likely to hamper achievement

of the cities' visions. To achieve these goals, we first identified various Open Data resources and their characteristics in each city (Section 4). Then we presented our findings based ten impact themes and identified gaps in city visions and relevant actions (Section 5), and finally discussed challenges and provided some solutions. As a result of this research, we have also produced a data catalogue²⁹ with links to the surveyed data resources and projects starting by Birmingham and the West Midlands region. It is a future direction of this work to produce similar data hub for the other five UK smart cities.

According to the Technology Strategy Board's Future Cities Programme [25], the majority of UK cities have developed similar solutions, generally focused on utilising Open Data platforms to improve access to information, and facilitating delivery of new, better or more efficient services, whilst responding to decreasing public sector budgets. However, our findings show contradiction between the availability of static Open Data resources and the cities' visions. Among ten major impact domains, which we identified, (Art & Culture, Built Environment, Business & Economy, Community, Energy, Education, Governance, Health & Well-being, Natural Environment and Transport & Mobility), only a few of them are properly linked to the innovation actions and projects in cities, and the majority of these innovative projects are only supported by the static Open Data available by the data hubs and government website. Dynamic data resources that are accessible through APIs can also be sources of innovation by enabling services providing or relying on up-to-date information. We believe that while the dynamic data resource can produce quick wins, static open datasets holding cumulative historic records are required for data-driven research yielding the deep insights required for fundamental transformation of cities.

We conclude that the UK Smart Cities dream will be only achieved in reality, where those involved in Smart Cities-related projects cooperate and share experiences and resources to avoid wasteful duplication of effort and repetition of the same mistakes, and to maximise progress towards the common goal.

Open Data represents a long-term investment. It involves considerable short-term costs in terms of collection, cleansing, storage and curation of data, with no guaranteed return. In-house teams and external enterprises and entrepreneurs utilise the data to generate information that can be monetised through innovative services benefiting their customers and the city as a whole. Attracting and nurturing such a service-providing ecosystem is likely to require active measures.

There is also a network effect in play the more compatible sources of data that are available, the more valuable each data source is. Positive return on investment may thus only be achieved once a critical

mass of data is accumulated. The order in which data is collected and published is therefore important starting with datasets whose value is not so dependent on the availability of other data, and adding other datasets as value multipliers later makes sense.

This network effect is also a strong reason for cities to cooperate publishing compatible and complementary datasets will increase the value for all. Sharing of experience and coordinating initiatives will also accelerate progress up the learning curve for all. We therefore suggest that new procurement models may be needed in the public sector to allow small, innovative companies across UK to participate in the innovative growth.

The existing real-time (dynamic) data can, of course, be captured cumulatively to fill in at least some of these gaps, but it must be cleaned, indexed and curated in such a way as to facilitate analysis in conjunction with other datasets. Comparing the types of innovation (and the associated benefits) enabled by real-time data and cumulative, historic could be a future direction of this research.

ACKNOWLEDGEMENTS

This research was undertaken as part of the Big Data Corridor: a new business economy, which is co-funded by the European Union under the European Regional Development Fund (Project 12R16P00220). The results presented in this work was purely conducted by the Birmingham City University research team at the time of the paper submission.

REFERENCES

- [1] ST/ESA/SER.A/352 (2014) World urbanization prospects: The 2014 revision, highlights. United Nations, New York.
- [2] PAS 181: 20014 (2014) Smart city framework guide to establishing strategies for smart cities and communities. BSI Standards Limited, UK.
- [3] BIS/13//1209 (2013) Smart cities background paper. Department for Business, Innovation and Skills, London.
- [4] Kehoe, M., Cosgrove, M., Gennaro, S., Harrison, C., Harthoorn, W., Hogan, J., Meegan, J., Nesbitt, P., and Peters, C. (2011) Smarter cities series: a foundation for understanding ibm smarter cities. *Redguides for Business Leaders*. IBM.
- [5] Falconer, G. and Mitchell, S. (2012) Smart city framework: a systematic process for enabling smart+ connected communities. Cisco Internet Business Solutions Group (IBSG), San Jose, USA.
- [6] Centre For Cities (2014) Smart cities. Centre For Cities, London.

²⁹<http://birminghamdata.org>