DATA FOR LONDON
A City Data Strategy

SUPPORTED BY
MAYOR OF LONDON
This document contains substantial material, concepts and findings extracted from the PhD thesis “Data as Infrastructure for Smart Cities” written by Dr. Larissa Romualdo Suzuki at UCL.

How to reference and cite the source material of this document:
URL: http://www0.cs.ucl.ac.uk/staff/l.romualdo/DataInfraForSmartCities/
CONTENTS

Introduction 5
Strategy Overview 8
    How we will measure impact? 14
Policy Context 15
City Data Market: High Level Framework 17
Strategy Theme 1: Developing London’s Data Market 18
    Where we need to be 18
    Priority Actions 19
Strategy Theme 2: Organising City Data for Impact 21
    Where we need to be 22
    Priority Actions 22
Strategy Theme 3: Recognising the Value of City Data 24
    Where we need to be 24
    Priority Actions 25
Strategy Theme 4: Building Public Acceptance 27
    Where we need to be 28
    Priority Actions 28
Strategy Theme 5: Active and Effective Governance 29
    Where we need to be 29
    Priority Actions 30
Strategy Theme 6: Technology Road Map 31
    Where we need to be 31
    Priority Actions 34
Introduction

A Vision

Welcome to the City Data Strategy for London. This document is built around the following vision:

We want London to have the most dynamic and productive City Data Market in the world. In our City Data Market, the capabilities, talents and capacity of all our city data partners will impact on our huge social, economic and service-based challenges. To make this happen, friction in the sharing and value-driven exploitation of city data will be reduced to a minimum. City data will be recognised as part of the capital’s infrastructure. We will use it to save money, incubate innovation and drive economic growth. And London will achieve global renown for data impact.

We will of course use this strategy in support of City Hall initiatives like the Smart London Plan, but we acknowledge at the outset that the Greater London Authority simply cannot deliver this strategy on its own. Indeed, data knows no boundaries and is hard to contain, so we should not try. Our simple aim is to make sure that London, its economy and its communities are able to derive maximum benefit from the undoubted potential London’s Data Market can deliver.

It therefore signals the start of a plan which will actively integrate and mobilise all the ‘working parts’ of the city data economy. We want the various audiences that the Strategy is aimed at to engage with it. We will listen and evolve the strategy as this rapidly developing part of the cityscape develops.

The Case for a City Data Strategy

Data offers new insights and ways of delivering better services to citizens - from public health and social care, to refuse collection, public transport, and to retail in the private sector. New approaches to managing already hard-pressed infrastructure are also possible and needed. On all fronts, we need to continue to make London’s growth story1 a positive one, and to keep the city running smoothly.

Today, the pace of technology change is staggering - we communicate constantly on social media and mobile devices and depend upon digital technology to bank, shop, access services and navigate our way through everyday life. Ninety per

---

1 GLA Demography Projections: www.data.london.gov.uk/demography/population-projections/
cent of the data circulating on the Internet today was created within the last 2 years\(^2\). And the emergence of data is not about to slow. Emerging sectors like the circular economy have the potential to generate new forms of data to which we must try to give structure, organisation and form, if they are to function properly. The city needs to be planned and built with data, and future data exploitation, in mind.

So our capacity to understand how people live and places work, and to create economic impact and social value from data and technology has never been stronger. In the tech economy, we have the modern economy’s poster child. The Centre for Economic and Business Research estimates that ‘big data’ alone will create 58,000 jobs over the period 2012-17\(^3\). There are 40,000 digital businesses and 200,000 employees in London’s technology sector\(^4\).

Yet data does not quite seem to flow into and benefit the economy and society as it should. We are not always able to access the data we need to answer city challenges. The technical architecture falls short of what is needed. Regulation and governance are more often than not found to be responding to, rather than proactively creating the bow wave for full data exploitation. Culturally and organisationally we do not seem quite able to connect up the new data economy skills with our city government challenges. As individuals, we often have little idea of how we are creating data and the uses to which it is put by various parties.

**Mild Improvement by gentle Increment: why Open Data is no longer enough**

While this city has a fine history of open data, there is a growing recognition that much more can be done to exploit the value of ‘city data’ – a new definition we want to firmly establish at the centre of this City Data Strategy. Sharing and using city data more effectively can deliver improved public services, more efficient infrastructure, and wider data-led innovation.

Data and the technologies it underpins now need to deliver social, economic and environmental outcomes that are tangible to politicians and the communities they serve. At its simplest, data has the potential to transform public services and urban policy, and drive innovation. City Government cannot waste this opportunity and it must bring together the new disciplines and specialisms in the emerging City Data Market to make this happen.

\(^3\) Centre for Economic and Business Research, Data Equity: Unlocking the Value of Big Data (2012)  
\(^4\) Oxford Economics commissioned by London & Partners, June 2015
A broader Open Culture is now needed

Making data the agent of change it truly can be and securing value from it needs three things to happen.

1. City Government needs to be much more deterministic in terms of:
   a. Better anticipating how data-driven service models can change the city economy, services and society.
   b. Identifying and harnessing the opportunities for London’s data prowess to provide answers to city challenges (e.g. improving health) and to develop the data-driven products and services which will make a reality of smart cities and the internet of things.
   c. Addressing the challenging commercial and societal challenges the data revolution throws up at us.
   d. Adopting clearer positions to deal with emerging and disruptive technologies like Blockchain and distributed ledgers, and semantic web technologies (e.g. linked data and ontologies).

2. Securing a reliable supply of valuable city data, through a sensible and collaborative approach to harmonisation and sharing of data in London’s public services and private sector.

3. Building greater public understanding and acceptance of how personal data can and should be used, securely and safely, to drive uptake of new data-driven services and products.

An Open City Hall needs your Help

Change is already starting to happen in London. Now we need to organise ourselves more strongly around achieving the potential that lies untapped in the data that remains stowed away, as well as in the talents we can group around it.

This document signals the need for further work in complex and multi-faceted technical, policy and regulatory areas. It is also concerned with equally difficult questions of culture and organisational capacity in an emerging field.

We are clear on two things. City hall does not have all the answers, and we are not able to reach the best solutions without adopting a collaborative and wholly open approach.

We therefore invite you to be part of what we believe to be one of the first exercises of its kind in a global city. Help us to make a reality of the City Data Market to the benefit of all - business, government, and citizen.
Strategy Overview

Core Principles

The London data landscape in 2016 is increasingly complex and crowded. A range of technical, strategic and organisational issues in the city data supply chain need to be addressed to realise the full potential of cross-domain city data. Among these are data interoperability, platform leadership, building enabling approaches which allow open and proprietary data providers to co-exist and co-operate, and bringing clearer definition to policies and regulations which will enable users to exploit new data-oriented business models.

This City Data Strategy is therefore built around the following core principles, all of which are designed to promote the value of city data.

- The city data economy will be based on simplicity and alignment for data stakeholders. The city data economy needs to be made less fragmented and complex for data creators (London’s public services, utilities, private sector, citizens). This will be achieved through promoting technical advances in data sharing, harmonisation, interoperability and
open standards, as well as tackling head on issues on pressing policy issues such as data privacy and security.

- **A clear ‘City Data Market’ sharing culture will be promoted** within both the public and private sector. This will transform current data management practices and develop a philosophy that emphasises the exploitation of city data across the whole range of city policy settings and activities.

- **Data infrastructure will be co-ordinated and joined-up** to facilitate deep knowledge discovery, and the creation of new valuable integrated services. The exploitation of rich, interoperable and engaging cross-domain city data is key to this.

- **Non-technological elements and technical domains will have equal status.** Business models, value networks, feedback loops, a data marketplace, data toolsets, clear licensing arrangements, and more efficient data governance, are all vital building blocks of a functioning city data economy.
The Six Strategy Themes

This strategy is structured around six broad themes:

- Developing London’s Data Market
- Organising City Data for Impact
- The Value of City Data is properly recognised
- Public Acceptance
- Active and Effective Governance
- Technology Road Map

Under each is a short statement on the current context and further statement of our desired position. For each theme we have identified a compact list of priority actions we will now take forward to ensure delivery of the strategy.
Who is this Strategy for?

This strategy is aimed at three core partner groups, all of whom fill distinct roles, and sometimes multiple functions, in the City Data Market.

- **Structural partners:** those who actively promote the strategy and the impacts of the city data infrastructure and exploitation. This group includes the GLA, London public services, private sector representatives, the Open Data Institute, Tech City UK, the London Grid for Learning, academia, a Council for Data Ethics, and regulatory and standards bodies.

- **Supporting partners:** the providers of services, tools and data, individuals and organisations involved in defining market needs and the early stages delivery of city data infrastructure, and validators of the wider strategy. This group includes open and proprietary data publishers as follows:
  - London public services: including London Boroughs, housing providers and health services.
  - Utilities: water, waste, & energy.
  - Telecoms: Mobile and broadband companies.
  - Private Sector: including small and medium enterprises, major and global businesses.
  - The wider GLA family: including the Metropolitan Police and Transport for London, and Mayoral Development Corporations.

- **Contributing partners:** also potentially including the first two groups, these are the direct users of the data infrastructure, creators of business cases, providers of feedback, and creators of knowledge and insights from city data. This group includes:
DATA FOR LONDON: *A City Data Strategy*

- Data enrichers: who cleanse, maintain, augment, and manage data.
- Data integrators: who access, integrate, analyse, and publish data.
- Data consumers: who use, reuse and exploit the full value of data.

**City Data can take a Variety of Forms**

A core element of the strategy is the definition of ‘city data’. This is a very deliberate move to build on the open data proposition. We want to identify uses for a wider supply of city data from a more diverse set of data ‘owners’. This should include not just government and ‘public services’, but private sector organisations (e.g. utilities, infrastructure providers, retailers, land developers), all of whom have a different range of considerations with regard to data and how it is shared and exploited. Therefore, for the purpose of this strategy, the following terms and definitions apply.

- **Open data**: non-privacy-restricted and non-confidential data. Produced with either public or private resource. Made available without any restrictions on its usage or distribution.

- **Private data**: restricted and/or licensed data including permission, privacy, publication and distribution; as well as data that is presently held privately merely because it has not as yet been recognised to offer value. Produced with either public or private resource.

- **Commercial data**: licensed data including permission, charging, use and distribution. Produced with either public or private resource.

- **Sensory data**: open and/or restricted data collected by sensors, actuators and devices owned by public and private sector, and citizens.

- **Crowd-sourced data**: data provided, collected and distributed by citizens through the use of digital technologies and social media.

**Making London Datastore the Home to Data-led Innovation**

We have no desire to achieve monopoly status for London Datastore in this regard and are acutely aware of the existence, and very high likelihood of other stores of particularly service-oriented data. However, the platform does already contain a comprehensive stock of datasets, blogs, code to be shared and examples of best practice. We therefore want the city data market to recognise it as the first port of call for data exchange, support, ideas generation and collaboration.

Around the themes in this strategy we will be establishing a series of **toolkits and other supporting materials**. These will help the various elements of the London City Data Market to develop and deliver the strategy and will hopefully build capacity in our own and other contributing organisations.

These can be found on the **City Data Strategy pages** of the London Datastore.
LONDON DATASTORE:
OPEN DATA INSTITUTE ‘OPEN DATA PUBLISHER OF THE YEAR – 2015’ AND HOME TO 40,000 UNIQUE VISITS PER MONTH

London Output Area Classification

Connectivity Map

Bromley Healthcare Community Appointments 2014/15

London Tube Station and Journey Stats
How we will measure impact?

We will ensure this strategy is aligned with London’s core strategic objectives and smart cities vision. We will continuously assess the progress and related outcomes of our strategy using the following critical success factors.

- **Institutional Capability**: how successful are we in increasing cooperation in knowledge creation, creating public and private sectors alliances to deliver data infrastructures and set standards and regulations? How have we opened the City Data Market to wider participation, to drive cooperative knowledge creation and innovation?

- **Social Capability**: have we encouraged the active participation of urban residents in city service improvement? How effective has our support to the development of community-led initiatives, educational institutions and city R&D operations been?

- **Innovation Capability**: what impact have we achieved through local experimentation, continuous open innovation, and innovation transfer?
Policy Context

This City Data Strategy is linked to, or supports, a wide range of city policy areas:

- **Developing and managing London’s infrastructure.** The London Infrastructure Plan is clear on the role data has to play in co-ordinating infrastructure delivery, managing assets more effectively and supporting utilities and infrastructure providers in complex business planning through tools like the London Infrastructure Mapping Application.

- **Energy Policy.** The recently published London Energy Plan sets out future scenarios for energy demand and means of supply. Modelling complex data from a range of sources is key to arriving at a better understanding of our energy future, as is the provision of data to households to allow them to manage their own energy usage. To this, other emerging policy areas like the sharing economy need to be supported by a stronger approach to data management.

- **The broader Smart Cities agenda.** The Future Smart London report puts city data, and the London Datastore, at the centre of smart city development. Without a coherent approach to urban platform development and a city data architecture which allows us to integrate data flows and take advantage of the Internet of Things and sensor technology, mobile data and social media, we risk delivering sub-optimal value and missing opportunities to achieve scale from our current crop of demonstrator programmes.

- **Strategic Decision Making at City-level.** Expertise in data and computational science across business and academia (e.g. CASA at UCL, the Data Science Institute at Imperial and organisations like the Alan Turing Institute) can still be better aligned with city policy and strategy development, and problem solving in complex city systems.

- **Better Public Services and Proper Digital Government.** From integration of services like health and social care, to route optimisation and behaviour change in waste management, data sharing and powerful analytics, supported by effective city technology infrastructure, are key enablers to service efficiency and cost savings. In a broader sense, we need to organise ourselves around efforts to use data in the city’s digital innovation ecosystem.

- **City Technology Infrastructure.** Efficiency and associated cost savings caused by the aggregation of IT, development of stronger asset registers, and adoption of cloud services aside, it is only through common platforms capable of managing the capital’s data that the true innovation potential of data and the economic benefits of smart cities and digital government can be realised.
Citizen Trust and Privacy. Making sure security and privacy issues are carefully considered will encourage citizens and households to trade the ever larger volumes of data they produce so that value can be created from it. Stronger policies will also encourage acceptance and uptake of the next wave of digitised, data-driven services and products.

City as a Platform. We need to make it easier to harness technological capabilities in London around city opportunities and innovation challenges. Releasing public data encourages developers to build services that are trusted by citizens, government, consumers and the private sector. City Hall should feedback to Londoners how it uses data they generate to improve services.
City Data Market: High Level Framework

This diagram shows how the three categories of partners in the City Data Market are grouped around the various elements of the city data infrastructure.
Strategy Theme 1: Developing London’s Data Market

We want to build and operate an efficient City Data Market and create supporting tools so that the value of city data can be fully exploited. This means making data available to create better city services, better functioning city infrastructure, and economic impact. It also means providing the means to change organisational culture and capacity so that open innovation is an accepted approach and one in which all parties can play a part.

Context

In the emerging data market, the logistical distribution of city data is deeply fragmented. Data is produced by a multitude of systems, devices and applications, but much of it remains hidden from view. This distribution varies according to suppliers and their sectors (citizens, the public and private sectors, utilities), distribution channels, and the policies and regulations to which data is subjected.

Cities as Platforms

To evolve, cities must be viewed as platforms, with populations encouraged to utilize technology to creatively disrupt and redefine core functionalities. Every digitally enabled citizen living in a city is a hub of real-time data. When analysed in isolation, there’s no actionable intelligence. But when you view the data we produce on a macro scale, the possibilities for radical inventiveness are endless.

Gerard Grech, Chief Executive, Tech City UK (Cities as Platforms, techcrunch)

Where we need to be

Better city data can lead to better city services, better urban policy and can spur innovation. To reach this point, city government needs to be both more deterministic and more co-ordinated about the
data it needs and the uses to which it wants it to be put. We need to be much clearer in setting out what data is needed to ensure progress and establish the conditions in which both public and proprietary data can be published, shared, managed and disseminated. As a complex city government, we can do more to set out the potential for data sharing and establish activities to encourage for its exploitation.

We need to deploy state of the art data management strategies alongside business and financial models, and policies and licensing agreements, to deliver a foundation for widespread exploitation of data over the long term.

We will do so around the concept of value. We will build a strong value-oriented network of partners to reduce complexity and costs associated with data infrastructure development, and ensure London’s data has the necessary quality for knowledge discovery and creation of new data driven services and smart infrastructure.

**Priority Actions**

We will:

1. **Assess, plan and build a pan-London Centre of Data Innovation.** This will seek to involve the wider public and private sector, and specifically the London Boroughs as significant providers of city services. Collaboration will be on either a project-specific or data fellowship basis (to aid the transfer of skills). We will also build capacity in allied from legal advice, human-centred service design and computational science. This effort will be tied to plans to develop a dynamic digital infrastructure for the capital - the London Supercloud. We will make the case for pan-London activity of this kind to the sponsoring departments for this agenda and the Government’s forthcoming Digital Strategy – the Department of Culture, Media and Sport, the Cabinet Office and HM Treasury.

2. **Make London Datastore a ‘convening hub’ for data-led innovation in the capital by**
   - Actively collaborating with public bodies and the private sector parties (e.g. Hypercat) that are integral to London’s data potential.
   - Delivering a series of practical features (e.g. payments, advanced search across different platforms) to enable data subscription models and financial transactions by providers and users.
   - Carrying through a clear plan for a series of technical enhancements which will enable the secure sharing of sensitive and new forms of city data, and which will be able to manage the flow of data to be generated by the Internet of Things.
   - Publishing a series of toolkits, relating to the main elements of this strategy and designed to provide practical advice for city data market partners.

3. **Work with London Boroughs to establish a further set of use cases for first, the sensible harmonisation of data at the point of sharing, and second, to prove the worth of data science approaches and digital innovation.** This builds on the GLA and London Councils’ successful application to the Cabinet Office’s Data Science Accelerator Programme.

4. **Establish datasets that are subject to greatest demand from the three different categories of city data partners (structural, supporting and contributing), and use them for the trial of standards and new technologies, and to explore and promote use cases where data is used to deliver different**
forms of value. This will require consultation (e.g. with the Mayor’s Infrastructure Delivery Board, London Councils, the Alan Turing Institute).

5. Build and manage a portfolio of initiatives that will help raise awareness of the potential of city data and change culture. This will include running at least two data discovery projects a year. These will be based on genuine policy need and will aim to establish city data analytics and the application computational science at the heart of public service improvement, city operations and wider innovation activity in the capital.

Developed by the GLA, the London Schools Atlas is an interactive mapping application, providing a detailed picture of London Schools, including performance, distance travelled by pupils, as well as current patterns of attendance and future demand for schools places.

Different levels of functionality and layering of data allow it to be used by schools place planners and policy makers in London Boroughs, as well as by parents exploring school choices for their children, including academies and free schools. Containing millions of individual pupil records, as well as time series and detailed geographical units, it presents a strong argument for harmonised data or better organised registers (the pupil level data comes from the National Pupil Dataset) for the 32 London Boroughs and can help provide a strategic and city-wide view of a public policy issue that cannot be strictly contained within administrative boundaries.
Strategy Theme 2: Organising City Data for Impact

Better organisation of City Data is needed if it is to have a strong underpinning role in public service decision making and infrastructure investment decisions, and broader innovation. It can be used more systematically and rigorously to manage city services and assets, and can also be more effective in creating a true ‘City as platform’ approach and new business models.

Context

Organising for the Internet of Things and Cloud
We have built the foundations to harness the power of the Internet of Things and the Cloud, through London Datastore making data a core part of London’s infrastructure. We must succeed in keeping London at the forefront in a data-driven world and invest more in London’s data infrastructure. Doing so will keep the city brilliantly placed to deploy the creative and game-changing power of technology.

Professor David Gann, Chair of the Smart London Board

Progress made in London on open data publishing has been impressive but is now in danger of delivering diminishing returns. A set of further technical, strategic and organisational issues in the city data supply chain make it difficult to capture the full potential of better organised cross-domain city data.

Data integration in a city setting is further hindered by the current lack of widely-accepted standards (e.g. for expressing the syntax and semantics of city data). A lack of understanding of the technology requirements introduced by the wide variety of data already available in the city environment also stands in the way of progress.

London (through the GLA) is the lead city in the European Commission’s Innovation Partnership ‘Urban Platform’ initiative to deliver impact in London and beyond. We are currently leading a group of European Cities in the development of a set of city-needs related requirements for urban platforms and the use of city data on them. These are freely available to all public services in London wishing to publish and exploit their city data.
**Where we need to be**

We want the London Datastore to provide a range of cross-domain data services that have a clear focus on using data to improve city services and meet societal and urban challenges. The platform should provide accessible services to users, and operate policies, regulations and standards which enable users to exploit new business models which deliver impact for the city. The platform should also have the capacity to handle open, private (commercial) and volunteered citizens’ data.

Standards have been adopted, adapted or developed and provide a framework for the organisation of data. Using the excellent work on the National Information Infrastructure as our starting point, ‘Core Reference Data’ will be created and maintained to allow the easy combining of data from different sources. Organisations from all sectors will be better able to analyse how things work now and how they can be made to work better in the future.

**Priority Actions**

We will:

1. Deliver a city data mapping exercise and a detailed picture of the architectural features of our existing city data, including: sources, volume, variety, temporal factors and sensitivity in order to identify the processes and resources needed to manage the city data supply chain (e.g. computer and network power to respond to data requests, data storage requirements).

2. Create viable (commercial) models with partners. These will enable useful and interoperable cross-domain city data infrastructure and consumption to develop. We will build on the work of the London Legacy Development Corporation and partners who are attempting to do this through the Queen Elizabeth Olympic Park Smart Sustainable District.

3. Work with contributing partners to make available a set of tools (e.g. data publication interfaces, extraction of descriptive and technical metadata, data quality verification mechanisms) to facilitate and automate the publication and management of data and meta-data in the London Datastore to ensure standards and licences are adhered to.

4. Offer open and well-documented APIs, as well as service descriptions and contracts, all of which will be available for reuse by other parties. This will enable developers and interested individuals to openly utilise the resources provided, and to encourage open innovation off the back of the City Data Market.

5. Make available the ‘Core Reference Data’ that underpins the ability to combine data from different sources, whether it is hospital names, sectors that companies operate in or geographical boundaries.

6. Carry out a review of existing metadata schemas (e.g. open data, sensory data ontologies). The purpose of this will be to identify those schemas with the greatest potential for reuse, to identify gaps, so that we can establish strategies to fill them.
7. Identify and publish integrated approaches to design and service delivery in the city data infrastructure so that we are able to ensure that appropriate standards are adopted (e.g. Standards Hub, IEEE P2413, W3C Semantic Web Standards).

**TRANSPORT FOR LONDON:**
**DEVELOPER ENGAGEMENT, BIG DATA, BETTER CUSTOMER EXPERIENCE**

The Unified API is TfL’s open data store. Live and reference data from operations is captured and transformed so they can be easily used by a large community of developers who generate new and innovative tools and services, without any direct investment from TfL. Forty per cent of Londoners now use mobile apps powered by data. We’ve already seen the impact of open data (6,000 developers, 460 apps, 200 API elements) and big data (automated refunds, re-planning bus network, better information) used to real practical effect. TfL continues to explore the potential for more datasets to be made freely and openly available to further boost these efforts.

The Oyster card was the first, and most visible, way in which TfL collected customer journey data. Analysis of it is vital in understanding transport users’ expected – and unexpected - journey patterns across the transport network.

TfL is now collecting and analysing data on customer journeys via more than the Oyster card. Since the launch in September 2014, there have been more than 265 million journeys made using Contactless Payment Cards. The proportion of pay as you go journeys being made using CPCs continue to go up each week, reaching 27 per cent for Tube and rail journeys and 23 per cent for bus journeys. CPCs can also be enabled on mobile phones and a number of such applications are now live in the market, including Apple Pay and mobile wallets run by Vodafone and EE. The launch of Apple Pay in the UK saw a sharp rise in the number of mobile transactions. Overall, in the second half of 2015 more than 3.2 million journeys were made using mobile devices. The share of mobile devices in overall contactless use has been growing and stood at approximately 3.5 per cent in December.
Strategy Theme 3:
Recognising the Value of City Data

The value of city data must be considered across social, environmental, economic, as well as and financial domains. By establishing better measures and monetising data, value-based transactions can be properly understood by all partners, and can therefore take place, in the City Data Market.

Context

Currently, the primary and secondary uses to which data can be put are poorly understood. This is as true for the private sector as it is for the public sector, but the latter is less likely to consider data as an asset with value attached to it. Because, amongst other issues, there is an assumption that value is inherent in data that people ‘own’, data sharing beyond the current open data approach, and the potential for value creation that comes with it, is limited.

There are some signs of the necessary cultural and organisational change. The London Infrastructure Mapping Application, delivered for the Mayor’s Infrastructure Delivery Board, features data from a range of utilities, and organisations outside of the public sector.

London’s data has social, economic and environmental, as well as financial value. Value across the former group is achieved when data-driven insights or product development by local public services, smart city applications and academic research and development activities are shown to tangibly improve outcomes for cities, their businesses and their residents.

Where we need to be

The maturation of smart city business models, future advances in the application of Internet of Things technology in the running of public services and the city in a broader sense, and the increased presence of data-driven businesses in the economy are already creating, or will create, new business models. We understand the importance of establishing a solid appreciation of value across all its forms, and the need for this to be based as monetisation, if these business models are to be successful.

The Greater London Authority wants to set a financial framework for the city data economy in order to define which investments, costs, risks and revenues are associated with the exploitation of both open and commercial city data.
Ultimately, we want city data infrastructure to support diverse business models for the open and commercial exploitation of city data. We think that charging users to consume data may increase competitiveness as data providers will be more inclined to provide high quality data to end users. In this way, data providers can also recover costs associated with the production and release of data.

We also realise that sometimes data ought to be exploited for purely social value but there is a cost to collation and management of the data. Understanding that cost will allow the social benefit to be better contextualised and allow strategic choices to be made.

**Priority Actions**

We will:

1. Establish a series of value cases for the exploitation of city data from public and private sources, setting out clearly in each how economic, social, environmental and financial value is delivered.

2. Commission an economic analysis to establish a monetisation framework for the City Data Market. This will allow the various categories of partners to understand the value they are contributing to and deriving from transactions in the City Data Market.

3. Engage with a range of existing and potential publishers of city data to establish the costs associated with data collection so that effective ways of recovering costs against the opening up of data can be explored (e.g. through investment sources and public and private sector partnerships).

4. Put in place the necessary enabling licencing agreements for open and commercial data to ensure we are maximising the potential for the use and re-use of individual city data sets (e.g. UK Government Licensing Framework (UKGLF), Open Government Licence (OGL)).
Through the Infrastructure Mapping Application, London has for the first time a detailed picture of the infrastructure development and how this relates to economic and population growth in London, and broader plans for urban development (e.g. housing and transport schemes).

Added to the GLA’s own data – demographic, economic, labour market projections – the platform contains approximately 8,000 data points from a range of infrastructure providers, and developers to deliver a range of benefits:

1) Utilities and infrastructure providers are better equipped to develop evidence-based business plans which meet London’s needs;
2) Developers are able to see with greater certainty the infrastructure pipeline, which helps unlock development sites more easily;
3) The Boroughs are able to form more cohesive local development plans; and;
4) Training providers are better able to respond to future industry requirements as they relate to infrastructure development.
Strategy Theme 4: Building Public Acceptance

The next era of data exploitation and the services it has the potential to create will only become a reality if security, privacy and trust issues in relation to personal data are adequately addressed. A greater appreciation on the part of the public about how their personal data can be used to generate value for the economy, for wider society as well as themselves, will lead to increased data supply and will ultimately drive public uptake of services.

Context

Citizens are an integral part of the City Data Market, responsible for a significant volume of London’s city data. In the main, this data mostly comes from everyday interactions with services, social media, mobile devices, vehicles, and smart cards. In the future, smart homes, other smart city services, and new forms of autonomous technology and artificial intelligence will only increase the amount of personal data generated. Often people have little idea about the ‘exhaust’ data they are creating as a result of their actions. Further, much personal data disappears into the data repositories of the companies who collect it, and its wider potential remains untapped.

Better Public Policy in the Big Data Era

Big data hold major promise for social science, which should enable us to further extend our record in policy research. We have access to a cornucopia of data of a kind which is more like that traditionally associated with so-called ‘hard’ science. Rather than being dependent on surveys, the traditional data staple of empirical social science, social media such as Wikipedia, Twitter, Facebook, and Google Search present us with the opportunity to scrape, generate, analyse and archive comparative data of unprecedented quantity.

Professor Helen Margetts, Director, Oxford Internet Institute (www.oii.ox.ac.uk)

Of course, the exploitation of volunteered citizen’s data raises new challenges in security and privacy, as well as consent issues. Well documented cases mean that there is currently a trust deficit when it comes to both the public and private sector’s ability to keep personal data secure and to always be clear about how it will be used, especially when data is being shared with third parties. Recent research by the Digital Catapult showed that 60 per cent of the British public are uncomfortable to share personal data, despite many doing so (maybe unknowingly) on numerous social media sites and internet-based services. Most tellingly, 38 per cent do not know what ‘personal data’ means, yet 96 per cent claim to.
Examples of sensitive data found in cities are transport historical data which contains user personal
details and transport usage patterns, surveillance data (videos and images), and smart metering data
(which might reveal information such as whether a household is occupied or not). And as a sign that
the use of analytics is not immune from controversy, ‘predictive policing’ has been criticised for
potentially leading to biased decisions in law enforcement and labelling some communities as suspects
of crimes not yet committed.

Where we need to be

A stronger mandate is needed. This means influencing and responding to a dynamic part of the city
data landscape.

We need to convince the ‘providers’ of personal data of the potential value in this data to themselves
and wider society, and to drive economic productivity. From combined personal data from different
sources allowing doctors to make personal health recommendations to patients; health data generated
from wearable devices allowing clinicians to analyse vast swathes of data in pursuit of medical and
scientific research; to households using more nuanced and detailed domestic energy usage data to
arrive at better deals from energy companies, there are numerous examples.

At the same time, people should implicitly expect and have confidence that data is secure and privacy
preserved. High levels of trust and privacy are a prerequisite for successful participation of the public
in a London Data Market that can genuinely be considered to be delivering cross-domain city data.

If this can be achieved, then the possibility of a new generation of responsive, tailored and waste
limiting services based on personal data can become a reality.

Priority Actions

We will:

1. Work with the Digital Catapult’s Personal Data and Trust Network, to promote responsible
innovation with personal data. We can offer to members of our Talk London community personal
data exchanges in which public and private sector use cases for personal data sharing can be
tested. This research will allow us to explore the suitability of the Trust Framework for the
management of personal data in the data market.

2. Establish a Consumer Data Council so that London citizens can be part of the discussions about
how their data can and should be published and used in the City Data Market. Its remit will cover
data and digital literacy, and engaging with a range of personal data protection and consumer
groups to ensure that Londoners better understand regulations around personal privacy rights.

3. Explore the privacy aspects of personal data, and define data management strategies for the
capital’s data infrastructure which ensure data integrity and compliance with national, European
and international data protection regulations.

4. Support calls by the Royal Statistical Society and the House of Commons Science and Technology
Select Committee for the Government to establish a Council of Data Ethics to address the legal
and ethical challenges associated with balancing privacy, data security and public benefit.
Strategy Theme 5: Active and Effective Governance

Active governance is now needed around London’s City Data Market so that strategic challenges can be addressed, and opportunities can be anticipated and capitalised upon. More prosaic but just as important, good governance will lead to a better understanding of regulatory and legal issues so that the sharing of data and benefit derived from it are both maximised, and harm limited. This activity is vital if new business models needed to underpin data-driven products and ‘smart’ initiatives are to emerge and succeed.

Context

The complexity of London’s City Data Market has increased significantly since the early days of open data. Governance around the London Datastore is currently ad hoc and concerned mainly with data quality and data release. There is little attention paid to issues such as data interoperability and harmonisation. There is also a set of broader questions to be addressed around the relevance of existing legislative and regulatory framework given the increase in the sources of city data, the further potential for connected devices to generate still more, and the growing range of uses to which this data can be put.

The current challenge is therefore one in which all of the complexity of managing a city data infrastructure and the potential for inefficiency, duplication and unreliability of data supply need to be minimised.

Where we need to be

Platform openness on its own is not enough to drive service improvement, innovation and culture change around issues like data sharing. Through strong governance arrangements, London can establish a leadership strategy which enables technology, data, and digital assets to be managed as city-wide resources. A key function of new governance arrangements must be to provide a context in which distributed innovation, driven by value creators in the City Data Market, can emerge.

To provide this, governance arrangements should oversee the ‘pipeline’ of data release in London. It should also be aware of the full range of potential data uses and be taking a positive stance to tackle practical issues that prevent data sharing and harmonisation, as well as regulatory and legal blockers.
In short, we need to know who wants to access the data, for what purpose and how this delivers social, economic or environmental impact in London. We also want to know what this data is (i.e. is it personal) and how it is being shared (i.e. in a safe and secure setting when this is needed).

**Priority Actions**

We will:

1. Strengthen significantly governance around city data by establishing a City Data Board made up of the key structural partners identified earlier in this document. The board will provide convening power, give strategic direction to the themes set out in this strategy, and govern the delivery of the actions relating to them. Reporting to the Smart London Board, the Board will also set the strategic direction for a London Centre of Data Innovation.

2. Set up a ‘City Data Exchange’ on the London Datastore to encourage all partners in the City Data Market to set out their data requirements. This would have the effect of achieving clarity on desired demand for data and the uses to which it is to be put, and will bring into the open issues around data sharing and harmonisation, value and monetisation.

3. Ask the London City Data Board to establish a series of data supply requirements across the range of public sector contracting arrangements and private sector business models, to prevent valuable data being locked up behind commercial arrangements and other considerations.

4. Through the City Data Board, push for broad access to data across government departments (e.g. HMRC and DWP) for London, to overcome issues which currently prevent further data sharing opportunities outside of departmental silos and which stand in the way of value creation.

5. Organise a sub-group of the City Data Board to offer professional guidance over individual decisions relating to data release and data sharing. This group’s goal will be to drive best practices for city data release and exploitation to ensure the quality, interoperability, trustworthiness and overall value of data offered and used. Its work will also align with that of any Council for Data Ethics formed on the basis of suggestions made by the Royal Statistical Society and other parties.
Strategy Theme 6: Technology Road Map

Effective co-ordination of the city’s Data Infrastructure through a technology road map needs to happen with urgency. If we fail, then London faces the prospect of missing the opportunity to provide true twenty-first century public services, and failing to integrate vital services like health and social care. We risk falling short in digital innovation and delivery, and generally limiting the city’s economic competitiveness.

Context

At the centre of this strategy is the need to provide the most advanced means possible for the sharing and exploitation of data by organisations and individuals. Much of the technology to support the vision and the actions in this strategy already exists in one form or another, and is planned for either through the GLA’s own programme of work, or beyond.

Within the GLA, the London Datastore is already using a cloud-based open source system (CKAN) and APIs to both harvest data from other data stores and present its own data.

Other programmes and projects put further technology developments into play, but these need to be assembled in a complementary way and tested at scale. There is further potential to customise this technology and make sure it is applied in further organisations operating in the City Data Market.

Where we need to be

We need to deliver, across a series of connected initiatives, the following attributes into London’s Data Infrastructure.

- **Sharing** - the sharing and presentation of information and functions to city data partners (citizens, city managers, other third-parties) through a variety of tools.

- **Sensing** – effectively providing processing and analytics capability to allow the refinement and consumption of data and information by the city or citizens.

- **Storage** - data storage is likely to be distributed, but catalogued so it is searchable and accessible from a single point.

- **Advanced Analytics** - the techniques deployed by data scientists to work with unstructured and big data require suitable software, hardware and storage to support them.
Cloud Technology for a Digital Innovation World Leader

Innovation, austerity, shared services, cross-sector working, devolution, smart cities, city-wide digital infrastructure, a breakthrough on big data, and the need for London to be seen as a digital leader, and the opportunity to do things once, or at best a couple of times, rather than 33 different ways across London are all compelling reasons to make the London Supercloud a reality. We are actively pursuing alliances across government, including with the Greater London Authority.

John Jackson, Chief Executive, London Grid for Learning

Where standards have emerged either in other cities, nationally or internationally and been proven, we will adopt them in the delivery of the strategy. These need to include technical, data or more general Smart City Standards. Where suitable standards do not exist but are needed to deliver the actions (for instance building up a London-wide dataset from each of the 32 Boroughs’ information), the GLA will take a lead through its links to data owners, national bodies and other cities.

Legacy systems prevail across London (not only in the public sector, but in many corporations). Our general approach will therefore be to create flexible interfaces (and customised arrangements where necessary) to minimise the changes that are required by data producers to their own systems. Over time as systems are replaced or upgraded these requirements will reduce.

Finally, we will also make links to initiatives in London’s public sector to consolidate existing IT activities to create a cloud-based digital infrastructure for the capital.
WITAN: A PARTNERSHIP WITH THE PRIVATE SECTOR TO DELIVER DATA INFRASTRUCTURE AND COMMERCIAL PRODUCTS

Working with Mastodon C – An SME expert in big data, we are building a next generation city modelling platform, funded initially by Innovate UK. For the first time, we are bringing together and making interoperable all of the models used in planning for the city’s future. Starting with demographic and economic models.

While Mastodon C will hopefully develop a commercially viable urban modelling platform, the GLA will have a secure environment for modellers and policy makers to explore scenarios and to store data that for commercial or privacy reasons cannot be open data. It will provide the basis for London’s equivalent of New York’s Data Bridge, (a system for secure sharing of data between agencies across the city).
Priority Actions

We will:

1. Create a secure data sharing environment for public services and other City Data Market Partners. Built through WITAN, the City Data Store will be an open source, cloud-based and cloud agnostic system, accessible through a web interface to address the needs of secure access by multiple users and periodic peaks in processing need. It will be based on an enterprise security model, which has role-based and organisation-based permissions.

2. Deliver an urban platform capable of handling the Internet of Things data at volume and speed. We will do so through London’s Horizon 2020 Smart Cities and Communities Programme ‘Sharing Cities’.

3. Help to build the London Supercloud (city-wide, cloud-based digital infrastructure) by aggregating demand across London public services for a range of big data functions, and a range of applications for digital innovation in public and wider services. To assist in the delivery of the improvements in public service performance possible, we will establish a formal partnership status with the London Grid for Learning, and other relevant groups like the London CIO Council.

4. Deliver a full and detailed set of functional and non-functional requirements for public sector bodies wishing to develop urban platform technologies. A draft version of these, created through our city lead role in the European Commission’s Innovation Partnership, has been created and is available here.

5. Define a modular based architecture to ensure that the London Datastore operates as one of a number of city data platforms and is able to accommodate additional functionality at a fair and transparent cost. This will rely on stable and well-defined interfaces, and will consider open source as an optional commercial model. As with all our work to this point, we will hold open standards as a central, guiding principle.
Other formats and languages
For a large print, Braille, disc, sign language video or audio-tape version of this document, please contact us at the address below:

Public Liaison Unit
Greater London Authority
City Hall
The Queen’s Walk
More London
London SE1 2AA

You will need to supply your name, your postal address and state the format and title of the publication you require.

If you would like a summary of this document in your language, please phone the number or contact us at the address above.

Chinese
如果需要您母語版本的此文件，
請致電以下號碼或與下列地址聯絡

Hindi
यदि आप इस दस्तावेज की प्रति अपनी
भाषा में चाहते हैं, तो कृपया निम्नलिखित
नंबर पर फोन करें और आवश्यक पत्र
पत्र पर संपर्क करें

Vietnamese
Nếu bạn muốn có văn bản tài liệu
này bằng ngôn ngữ của mình, hãy
hiến hột theo số điện thoại hoặc địa
chỉ dưới đây.

Bengali
অপরিহার্য অনুবাদকের ভাষায় এই ধারনার প্রতিলিপি
(কপি) চান, তা হল নীচের যে নম্বর নথিতে
কারণ বাংলায় অনুবাদ করে যোগাযোগ করন।

Greek
Αν θέλετε να αποκτήσετε αντίγραφο του παρόντος
εγγράφου στη δική σας γλώσσα, παρακαλείστε να
επικοινωνήσετε τηλεφωνικά στον αριθμό αυτού ή ταχυ-
δρομικά στην παρακάτω διεύθυνση.

Urdu
اگر آپ اس دستاوارگی کی نقل اینی زبان می
یابنچی ہیں، تو براہ کرم نئی کئی گھی نمبر
پر فون کرس یا دیکھی گئی پتے پر روابط کرس

Turkish
Bu belgenin kendi dilinizde
hazırlanmış bir nüshası
olunmak için, lütfen aşağıdaki
telefon numarası arayınız
veya adresi başvurunuz.

Punjabi
ਨੇ ਤਰੁੱਕੀ ਹਿੱਸਾ ਵਾਲੀਆਂ ਕੰਮਕਰਤਾਂ ਨੇ ਤੁਹਾਡੀ ਪ੍ਰਵਾਸੀ ਉਪਾਧੀ
ਦਿਤ ਕਰਵਾਈ ਨਹੀਂ, ਤਾਂ ਦੇਖਦੇ ਤੇਕ ਦੇਖਦੇ ਨਹੀਂ, ਦੇਖਦੇ
ਦੁਆਨੀ ਦਿੱਤੀ ਨਹੀਂ, ਤਾਂ ਫੋਨ ਕਰਦੇ ਹੋਣ।

Arabic
إذا أردت نسخة من هذا الوثيقة بلغتك، يرجى
الاتصال بقم الهاتف أو مراسلة العنوان
تلقائياً

Gujarati
જે તમારા નાની હર તસવીરો નહીં નથી પાછી ભાષામાં
સ્કીમની તોણ ન સ્કીમની તો, કેટલાક અપરિહાર
યુનન ઉપર
કોન કરવો આપવા નીચેના સરનામે સંખ્યા સાથે.