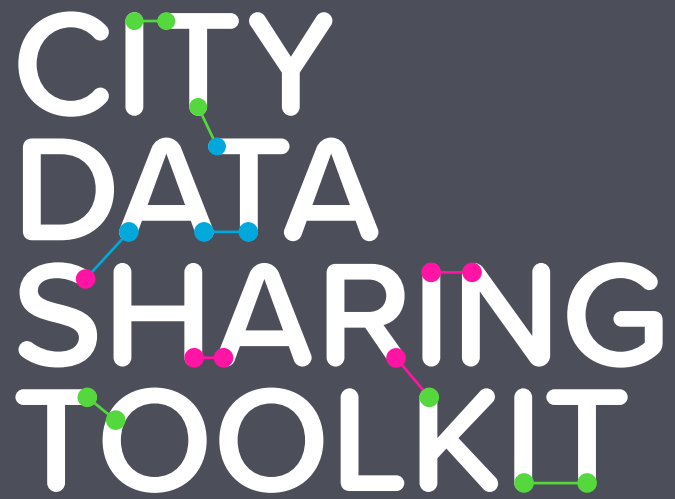# CITY DATA SHARING TOOLKIT

Tools, approaches and resources for city Governments sharing non-personal data

Version 1.1

CATAPULT Future Cities

This Toolkit has been designed to help people share non-personal city data across the public sector
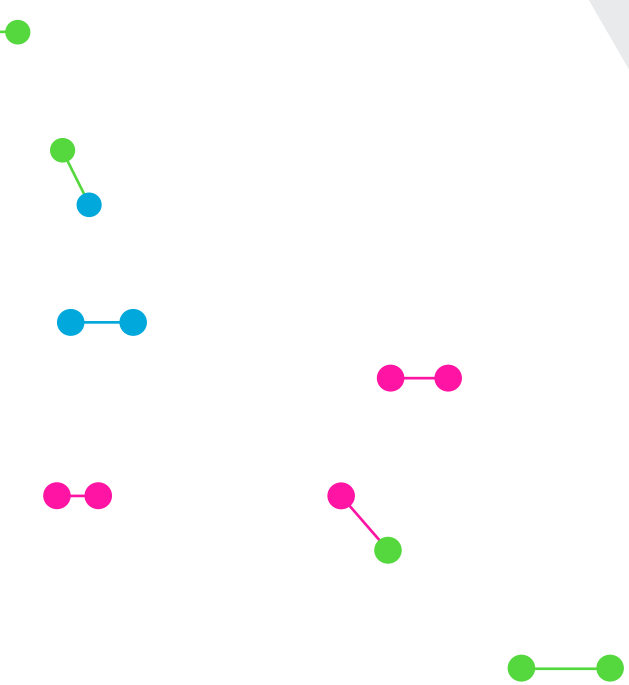
This toolkit has been developed by Future Cities Catapult to help organisations share city data. It sets out a high-level process for sharing data and includes links to further guidance and resources. For those working in public sector organisations, or closely with them, this toolkit provides guidance on setting up data sharing projects/processes.

As well as providing a high-level framework for how to develop your data sharing projects/processes, this toolkit also directs you to sources of further relevant information.

The toolkit deals specifically with sharing data, either internally or externally - we have assumed your organisation already has processes in place for the collection, analysis and use of data.

Based on detailed insights gathered from city data sharing projects around the UK, the toolkit was then refined at a co-creation workshop attended by a cross-section of professionals working across the public sector.

This is the first version of the toolkit and we are seeking to test and develop its contents in the cities we are working with. We would like this toolkit to become a resource for a wider community and have made it available **under a Creative Commons License so that anyone can adapt**, amend and re-publish this for non-commercial purposes.

CONTENTS

KEY TO RESOURCES

The toolkit draws extensively from existing resources, links to which (where possible) we have provided throughout this document. These resources include:

CASE STUDY
A short description of developing practice which illustrates a specific point

GUIDANCE
Policy and best practice produced by key organisations

REPORT
Think tank or support organisation report on relevant topic

BLOG
Personal perspective on a specific topic, sometimes offering useful 'hints and tips'

ONLINE LEARNING
Structured course or learning materials

TOOL
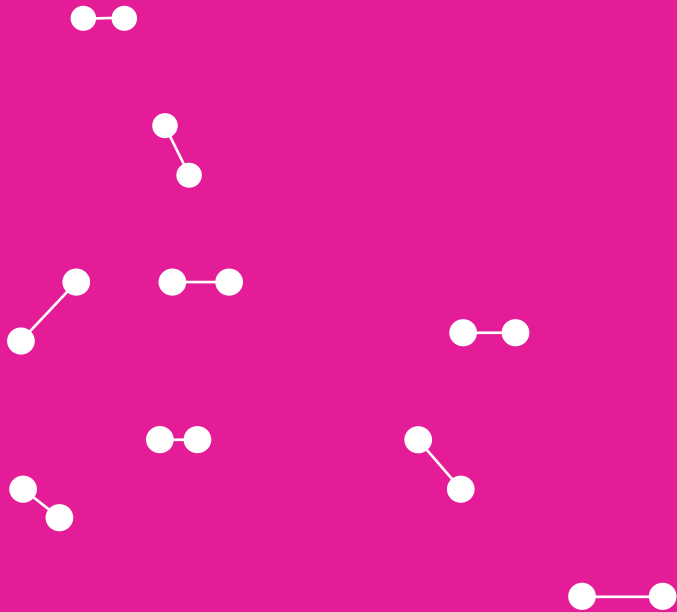Artefact or approach designed to help carry out a particular task

TOOLKIT
A collection of tools curated on a particular topic

Key resources are highlighted in bold at the top of each section.

BEFORE YOU START

WHY SHARE CITY DATA?

DRIVERS FOR DATA SHARING IN CITIES

There are a number of reasons for sharing city data. These include the growing amount of data, the need to collaborate to deliver services to meet complex needs and public sector digital transformation.

Increasing amounts of city data

New technologies and data approaches will allow cities to design the complex systems which will increasingly ensure cities are designed and run in the interests of citizens.

For instance, sensors are more frequently being installed by the public sector to monitor traffic flow in cities. Relevant data is also collected by driving apps on mobile phones and through fleet management systems installed by outsourced service providers (e.g. bin collection).

Cities are combining all this data along with public transport and even weather data to build more accurate and useful models of traffic flows in cities. All of this data will need an infrastructure of sharing in order for the best outcomes to be achieved.

Collaboration is inevitable

Shifting demographic patterns, the continued increase of populations in cities, climate pressures and constrained public finances all point to a greater need for collaboration between organisations. For instance, public sector organisations will increasingly look to collaborate with other parts of the public sector (and beyond) with similar aims to deliver outcomes to citizens. We are beginning to see these collaborations in the form of Combined Authorities, Local Enterprise Partnerships and Sustainability and Transformation Partnerships between NHS and local councils.

Joined up and personalised services will also require public sector organisations which have traditionally worked in silos to collaborate with each other, especially in terms of preventative measures where public health, sustainability and criminal justice are concerned. All of this collaboration will need greater data sharing.

Public sector transformation requires data

The public sector is also being encouraged to undertake service transformation to incorporate digital technologies and approaches. The example has been set by the Government Digital Service and the development of its Service Standard. Public sector teams around the country are continuing to develop their services so that they are increasingly available online and across a range of devices. Agile software development approaches are being used to ensure that services remain up-to-date and continually focused on user needs. These approaches require data of all sorts to help build these digital services and, once delivered, they will also create data on an ongoing basis.

DATA SHARING IS MORE THAN A TECHNICAL UNDERTAKING

Thinking about data sharing from a technical perspective is only half the picture. Data will only be shared between organisations when you have identified organisational synergies and fostered a culture of sharing more broadly. In addition, data sharing arrangements will only be sustainable if all of those involved recognise, and are happy with, the type of value exchange developed.

Sharing data drives collaboration

Data sharing, although it might first appear to be a technical undertaking, is also about building organisational collaboration. The aim of data sharing is to enable the sharing of information, knowledge and wisdom between and through organisations so that they are able to deliver some form of social or financial benefit.

We recognise that sharing of data is not always second nature for public sector organisations. In the short term, the same pressures that will ultimately force organisational collaboration - stretched finances, complex problems etc. - are also likely to reduce engagement as organisations initially attempt to deliver purely on the basis of their own resources. Sharing and collaboration may not come naturally and may be difficult. Organisations will need to see more clearly how the first data sharing arrangements help them achieve their own organisational aims.

In delivering yet more complicated services in potentially more complicated organisational relationships we will need to continue to focus on making sure that cultures align and that behaviours and approaches don't introduce barriers for sharing. In some instances this non-technical work that supports data sharing is likely to be more time consuming, impactful and potentially more complex than some of the technical aspects of conveying data from one organisation to another.

Sharing data shares value

Sharing data also necessarily implies some form of exchange of value - be that insight, resource or goodwill. By sharing data with another organisation you may expect to gain:

- value internally from the new insight
- financial resource for selling data to other organisations which then supports the delivery of your public task as an organisation.

Throughout this document we have highlighted where there are specific considerations dependent on whether you have provided or received data using these icons:

DATA PROVIDER

DATA RECEIVER
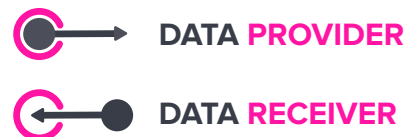
SHARING NON-PERSONAL CITY DATA

WHAT TYPES OF DATA DOES THIS TOOLKIT COVER?

There are many types of data which are collected by a range of organisations in cities which could include:

- Information collected from sensors
- Geospatial information
- Service provision and quality data e.g. transport timetables and real-time information, opening hours of services etc.
- Registers and lists of physical assets e.g. ownership of buildings, IT infrastructure documentation etc.

The diagram highlights a number of use cases within a city where data could be shared and we've highlighted some of the specific types of data which might be shared in each instance.

Primarily this toolkit is aimed at describing a process for cities' sharing of non-personal data internally within their organisation or between them and other organisations. By non-personal data we mean data which does not include personally identifiable information about a living individual (as defined under GDPR 2016 / DPA 2018).

It is important that all organisations handle personal data appropriately. Organisations are likely to be updating their systems and processes for handling personal data given their strengthened responsibilities under the UK Data Protection Act (2018) and the EU General Data

Protection Regulation (2016). We recognise that in some use cases in city contexts (including those highlighted in the diagram) the types of data which will need to be shared will include personal data. In these cases, the personal data which is identified will need to be treated in an appropriate way according to the arrangements agreed by the organisations involved and which respond to their new responsibilities under the DPA and GDPR. In particular, organisations will need to establish a lawful basis as described in Article 6 of the GDPR before processing personal data. The ICO produces a Code of Practice on data sharing in relation to personal data and will be releasing an updated version of this soon.

Throughout this toolkit we point to processes which cities will want to go through to ensure that the data they are sharing is non-personal in nature. However, many of the processes and approaches for sharing personal data are also helpful in handling non-personal data within organisations. Particularly useful are the governance and audit arrangements which have been set up in response to these new responsibilities.

Protection of vulnerable people

Potential data shared:

- Vulnerable people and their needs
- NHS data
- Police data
- Service delivery data

Better personalisation of social care

Potential data shared:

- Vulnerable people and their needs
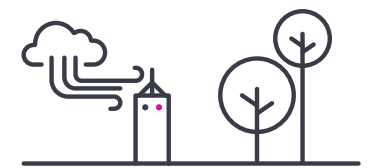- Addresses
- Sensor data

Optimising flows around the city

Potential data shared:

- Traffic light sensors
- Public transport data
- Mapping data

Planning and BIM

Potential data shared:

- Address data
- Mapping data
- Data about developers
- Utility data

Monitoring air quality

Potential data shared:

- Mapping data
- Air quality sensor data

8

WHAT DO WE MEAN BY SHARING?

'Sharing' can mean both sharing within an organisation and between organisations. It does not however cover public sector organisations and citizens sharing data between each other, or the sharing of city data between private sector organisations. In the majority of cases cited in this document we are referring to routine and ongoing sharing of data, but much of the advice in this document also relates to one-off sharing of a dataset.

COULDN'T ALL NON-PERSONAL DATA BE PUBLISHED AS OPEN DATA?

In short "no". Some shared data could be considered organisationally or commercially sensitive and although useful to be shared with other organisations it would not be appropriate to make it available without some control on access. There are also security and privacy reasons to restrict access to some data through a sharing relationship.

In addition, making data openly available may not capture the value exchange that some organisations would require. Some data is shared with a specific expectation of a reciprocal return of value, in particular in some cases data is sold.

For instance, in the case of geospatial data, some data is available as open data but other data products are available on a paid for basis. This recognises the value that those organisations get from using the data and returns value to the mapping organisation so that it can continue to fund the production of high quality data. However, Future Cities Catapult supports the publication of open data wherever possible.

SHARING BUSINESS RATES DATA IN BELFAST

Future Cities Catapult have been involved in a number of projects where city data has been shared. One of these was the Future Belfast project in 2016. We set up a collaboration with Belfast City Council to help them identify further revenue from business rates. In Belfast, business rates provide more than half of the Council's annual revenue which is used to fund services from waste collections to leisure facilities.

To develop, constrain and communicate the challenge, Future Cities Catapult worked with Belfast City Council to understand the rates process, the actors involved and the data landscape. As the project developed, it became clear that a data-informed approach was likely to be valuable and we began the discover, prepare and share process.

As part of this, we identified data relating to:

- Addresses
- Flood Inspections
- Property Licensing
- Commercial Waste Collections
- Building Waste
- Building Control
- Utilities

This data came from a range of sources, in a range of formats and with different levels of quality. It was also made available at different levels of granularity.

To make sharing easier and to enable fair access to all collaborators, the Council generated individual collaboration agreements with each partner. These collaboration agreements confirmed governance arrangements and addressed legal, privacy and ethical issues.

To initiate the sharing, all data was uploaded to a web-based shared space requiring credentials for access. This space also recorded the time and nature of any access requests for security purposes. If no longer using the data for the prescribed project, collaborators were required to remove the data from their systems.

IS YOUR ORGANISATION READY TO SHARE?

Before starting any data sharing project it's best to ensure you have the appropriate organisational data maturity.

This involves identifying if your culture, processes and systems are ready to commit sustainably to ongoing data sharing. Data sharing is more than simply the transfer of data from one organisation to another. As well as the systems needed to receive the data appropriately and store it securely in all of the organisations involved, there also need to be individuals in each organisation with the right skills to be able to collect, store and analyse the data.

There are a number of other frameworks and assessments in the pipeline. Nesta and the Local Government Association for example, have both developed online tools to support local public sector organisations assess their data maturity.

Our recently published Quick Start Guide to City Data sets out a framework to think about your city data project and an approach to thinking about your organisational preparedness for using data that covers Technology, Mindset, Resources, Skills, Operations and Legal issues.

TECHNOLOGY
MINDSET
LEGAL
RESOURCES
OPERATIONS
SKILLS
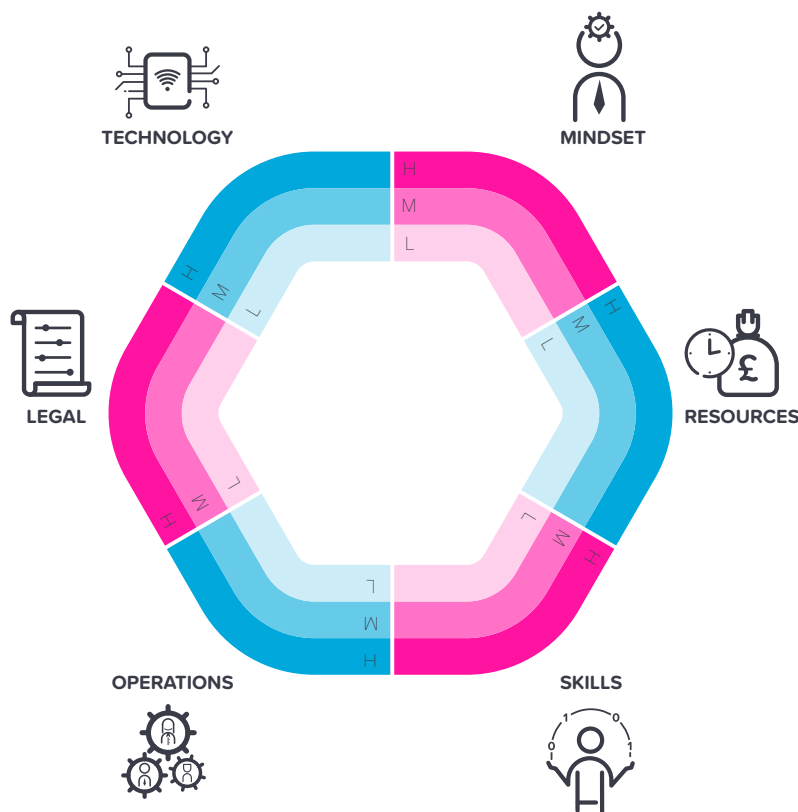
H - High   M - Medium   L - Low

In 2017, **DataKind** and **Data Orchard** (a data research consultancy) published **Data Evolution**, a report detailing research undertaken with social and charity organisations which sets out a Social Sector Data Maturity Framework. The report makes clear that the most important factor when it comes to levels of data maturity is in fact people, as opposed to systems and procedures.

Quick Start Guide to City Data (Future Cities Catapult) **- examples of challenges and opportunities for innovation in city data, as well as a loose framework for writing a City Data strategy as part of a city or region's wider data or digital strategy.**

Online organisational data maturity assessment (Nesta / LGA) **- an online tool for UK Local Authorities to assess the level of their current organisational data maturity and to benchmark this against their peers.**

Open Data Maturity Model (The Open Data Institute) - although this maturity model focuses on open data publication, there may be elements useful for public sector organisations thinking about the full data value chain.
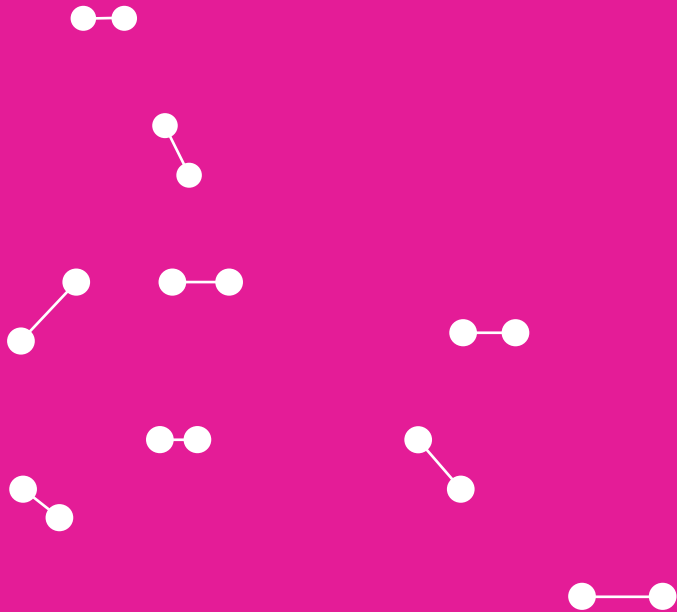
Data Fundamentals (School of Data) - an online course that provides "a solid overview over the workflow with data guiding you from what data is, to how to make your data tell a story".
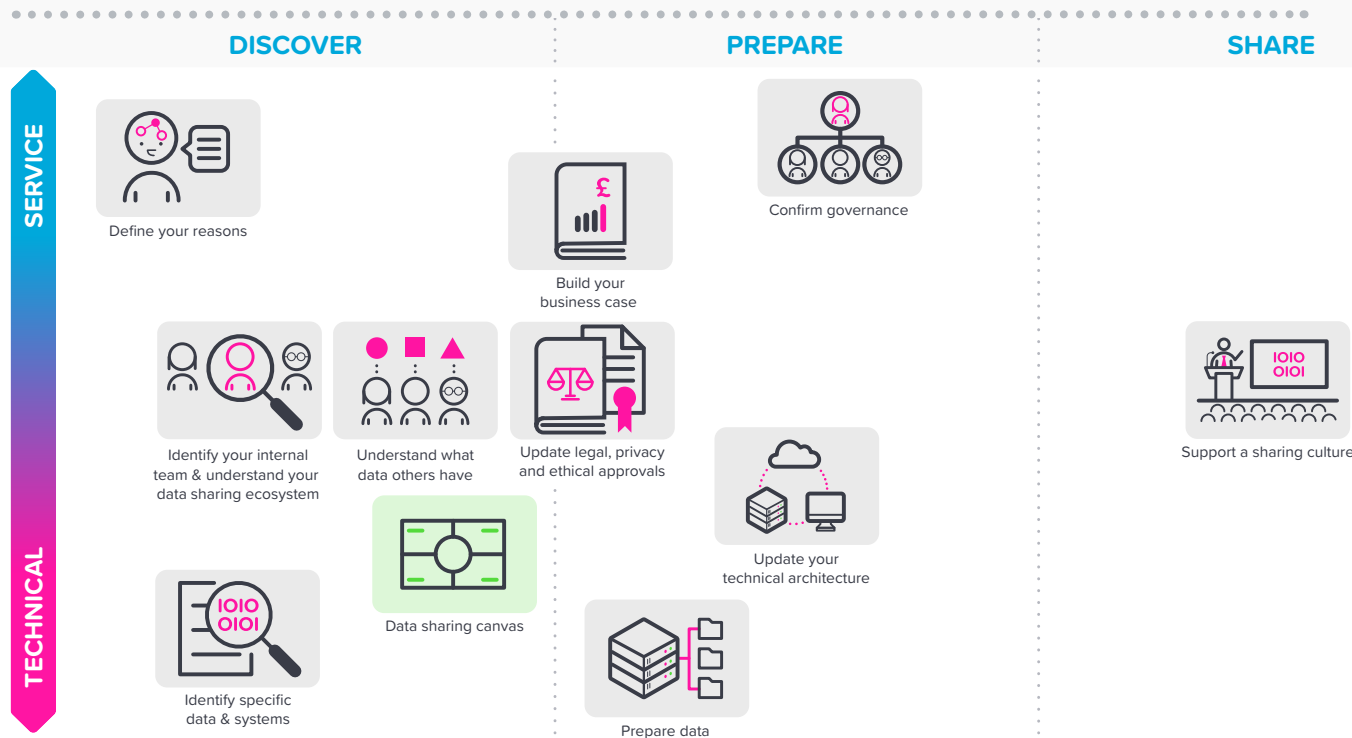
Solving Public Problems with Data (GovLab) - eLearning course focused more on data science and data analytical thinking for those wanting to work in the public interest.

A CITY DATA SHARING PROCESS

OVERVIEW OF THE PROCESS

DISCOVER

PREPARE

SHARE

SERVICE

Define your reasons

Identify your internal team & understand your data sharing ecosystem

Understand what data others have

Build your business case

Confirm governance

TECHNICAL

Identify specific data & systems

Data sharing canvas

Update legal, privacy and ethical approvals

Update your technical architecture

Prepare data

Support a sharing culture

The visual above sets out our initial thoughts on a city sharing data process in the form of a Kanban board. This Kanban board details those phases and jobs that we identified as core to setting up a city data sharing arrangement, and positions them according to who is likely to be leading on each aspect.

Each card on the board relates to a following section of the Toolkit.

NOT A LINEAR PROCESS

This toolkit was drawn from discussions with those who have undertaken data sharing projects in cities. Whilst it became clear that the process for making data sharing arrangements wasn't always straightforward or linear, we were able to define a number of general phases.

Please feel free to jump to the sections which are relevant to your circumstance or use in a different order.

A COLLABORATION BETWEEN TECH AND SERVICE

It also became clear that data sharing projects require both technical and service skills. Often the work is a collaboration between those with ownership of particular service issues needing data to be shared, and those with the technical responsibility and knowhow to set up the data sharing arrangements.

We have highlighted above where the skills needed at different stages are more technical or service (or both).

PAS 183 - Smart cities – Guide to establishing a decision-making framework for sharing data and information services (BSI) - guidance on establishing a decision-making framework for sharing data and information services in smart cities. For decision-makers in smart cities from the public, private and third sectors, it will also be useful to any city organizations wishing to share data.

Data sharing code of practice (ICO) - statutory code of practice developed to support the DPA 1998 and currently being updated for the DPA (2018). Practices included mostly relate to personal data sharing.

Information Sharing Toolkit Scotland (The Scottish Government) - developed by the Scottish Government with a focus on the sharing of personal data with some tools and approaches that may be useful. Practices included mostly relate to personal data sharing.

A CITY DATA
SHARING PROCESS

DISCOVER
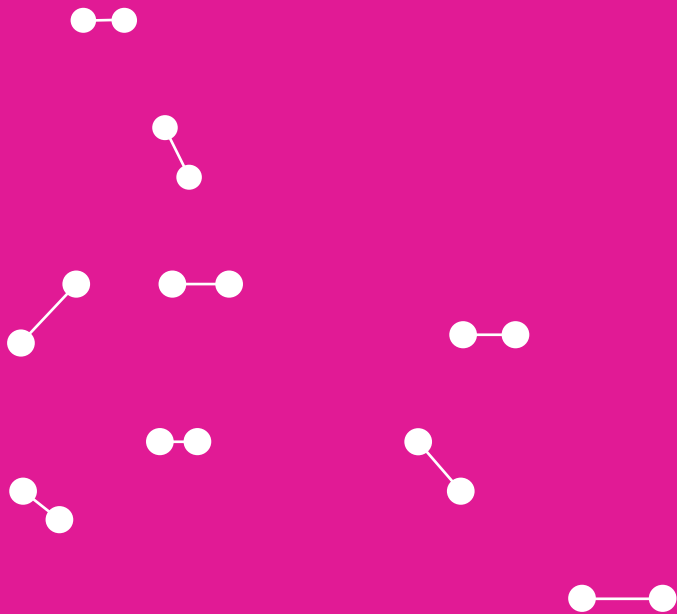
DISCOVER
DEFINE YOUR REASONS

Data sharing begins ideally with a clearly articulated rationale for sharing a specific data set. Clarifying this rationale, and understanding how it fits your business aims, helps build a strong supporting 'outline business case'. If necessary, this can be expanded later to a full business case (see later).

USER NEEDS AND ORGANISATIONAL AIMS

It is less likely to be the case that you have a statutory duty to share non-personal data. Therefore, you may have to go through a proactive process to identify the reasons to share data. Perhaps you have identified the need for a particular data flow through a service redesign process, or maybe you've outsourced particular services which now requires the exchange of data.

It might be useful to be clear how the data share will help you fulfil specific user needs and organisational aims. How will this data help a user fulfil their particular aim? How does this data share help the organisation achieve one of its aims?

SOME REASONS TO SHARE DATA

In our interviews with data sharing experts for this toolkit we identified a number of organisational reasons to share data:
☐ Gain greater insight into user needs
☐ Build a business case for your service
☐ Make decisions about individuals
☐ Help run a service more efficiently
☐ Bring in revenue
☐ Measure effectiveness of interventions / services
☐ Ensure regulatory compliance / meet incoming statutory obligations
☐ Ensure security of services
☐ Support procurement of service provision
☐ Identify and reduce duplication of provision

CASE CO-CREATION AND COMMUNICATION

Their early involvement in co-creating the project's 'outline business case' is a good way to build understanding and get buy-in. It will also ensure the project is clearly defined and, with all the key stakeholders 'on the same page', engagement and involvement with the data share's broader ecosystem (whether data providers or receivers) will be optimised. Suggestions:
• Identifying examples of similar projects could help identify benefits.
• Link projects to national or regional priorities could help make them more relevant.

DATA MINIMISATION

In the first instance you need to make the case that you're not able to fulfil the user need/ organisational aim without sharing data. One of the core principles of the GDPR is 'data minimisation' i.e. reducing the data collected/ held to only that which is necessary to perform your functions. It is good practice to consider whether you could achieve your aims without sharing data or through using other ways of achieving the insight you need. And if you do need to share data what is the minimum amount that needs to be shared and still remains useful for the purposes identified.

OTHER CONSIDERATIONS

In building this elaboration of your reasons for sharing you may also want to consider:
• **PRECEDENTS** - other examples of similar projects could also help identify benefits.
• **POLICY** - Link this current project to a national or regional priority could help make the work more relevant.
• **MATURITY** - What stage of development are you at in terms of the delivery of your service?
• **ANALYSES** - How is the data you want to share going to be used? Is it for strategic reason or operational?
• **SPECIFIC USERS OF DATA** - Who is going to use the data?
• **FREQUENCY** - How frequently will you need the data

DATA SHARING 'PATTERNS OF EXCHANGE'

As well as identifying the project's benefit to your own organisation, it's worth also defining the project's 'pattern of exchange' - how it carries benefit across all the organisations involved. This might be less essential where the data share is purely internal and to be expected. This can often seem implicit, but it's useful to be clear from the outset. On the basis of our investigations so far, these appear to be:
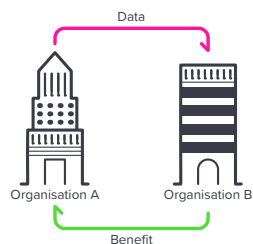
1. Mutual data exchange

Where both organisations have a need for data, the exchange creates mutual benefit.
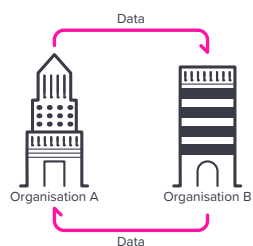
2. Data exchanged for other benefit

Where access to the data is exchanged for another form of benefit (e.g. financial, resource etc.)

3. Data exchanged for less specified/longer term benefit

Data sharing can be part of a broader organisational decision to make data more freely available with no immediate expectation of reciprocation. The benefits here are longer term and frequently less measurable. But can include greater collaboration within and across organisations that can in the longer run support a healthier environment for your organisational aims. This is frequently the argument for open data publication.
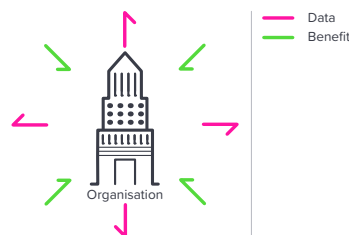
Here we've represented these relationships as only between two parties but clearly where there are more organisations involved then these patterns are likely to be more complex.

As part of the EU funded 'Sharing Cities' programme (of which Future Cities Catapult is a partner), the Royal Borough of Greenwich sought to define the most useful project to run in the programme's theme of 'eMobility'. Through engagement with the council leads on transport and parking, this was narrowed down. The user-needs identified were to enable electric vehicle users to find appropriate charging spaces, and for the parking enforcement team to prioritise their activities. This information is being shared through a secured API with Future Cities Catapult for use in data visualisations.

Urban Innovation Toolkit **- developed by Future Cities Catapult and Umbrellium, this online toolkit helps plan urban technology projects by methodically organising and joining up problems, stakeholders, methods, evidence and impact.**

Citizen Sensing Toolkit (Making Sense) - featuring lots of tools, this details a step-by-step approach to undertaking a citizen sensing project - i.e. working with the community to use IoT technology and analyse the data collected.

Public Sector Data Analytics - A Nesta Guide (Nesta) -comprehensive slide pack elaborating stages of a public sector data analytics project.

Sharing the benefits - How to use data effectively in the public sector (Reform) - recent report on the benefits of sharing data in the public sector.

Using open data for public services (ODI) - sets out a number of case studies on how open data is used in and benefits the delivery of public services.

Service Manual (HMG) - helps government teams create and run digital services that meet the Digital Service Standard.

Development Impact and You Toolkit (Nesta) - primarily for designing and running development projects, this could help clarify reasons for sharing from the outset of your project.

Data for Public Benefit (Involve) - This project by Involve focuses on elaborating the benefits of sharing personal data but may also be useful reading in relation to non-personal data

DISCOVER
IDENTIFY INTERNAL TEAM

Whilst you're in the Discovery Stage of the data sharing project you will want to identify those within the organisation who will be able to help you. We have tried to represent here some of the types of people you may want to engage in the process. Those we consider key are highlighted in bold. Not every organisation is the same and therefore some of these roles might not make sense in your particular circumstances.
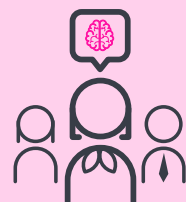
SENIOR MANAGEMENT

Chief Executive
• sets the organisational vision that informs all projects / services
• will want direct or clearly delegated oversight of final data sharing agreement

Elected Leader
• in very high-profile (in all likelihood rare) cases an Elected Leader or Cabinet Member may have a particular interest most likely when project forms a key part of a council's agenda

SERVICE DELIVERY

Head of Service
• **department/service head who would benefit from the data share e.g Head of Waste, Head of Children's Services etc.**
• **will help shape purpose of sharing data**
• **(you might be this person, or be in their team)**

Frontline Officer
• **crucial role in clarifying the reason for your data sharing project**
• **may understand why data unavailable in the first place**

ANALYSIS

Head of Analysis
• could clarify what data currently held by the organisation
• may have knowledge about data held by other organisations

Data Analysts
• **potentially many different analysts in your organisation support specific delivery teams or oversee performance reporting to senior management**
• **possess valuable insight about data quality and availability**

GIS Specialist
• Geographic Information System (GIS) specialists understand what data is held especially in relation to creation data
• they will also understand the capability of current systems

Data Scientist
• usually found in organisations where data used routinely and with greater data maturity (e.g. may use a programming language to automatically produce analysis of live data)
• likely to be someone who would use the data once it's shared

17

INFORMATION TECHNOLOGY
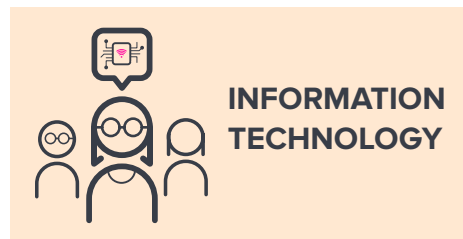
Chief Digital / Data / Information Officer
• may be a member of the senior team
• ultimately responsible for the use of data in the organisation

Head of IT
• will input on technical changes required when sharing data

Data Provision Lead
• **day-to-day contact for accessing data held in systems**
• **probably part of broader team generally responsible for data systems**
• **intimate knowledge about status of data, including its quality and comprehensiveness**
• **worth identifying and involving this person early on to understand as much as possible the reality of the data that might be shared**

Business / Systems Analysts
• intimate knowledge of the specific technical systems and their capabilities
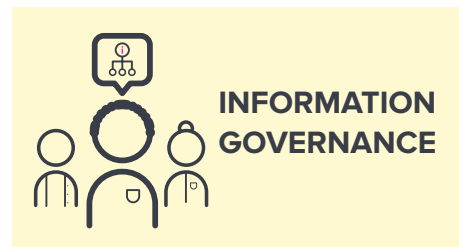
Technical Security
• probably in the IT team
• responsible for ensuring security of all technical systems

INFORMATION GOVERNANCE

Data Protection Officer
• formal role set out in the GDPR and DPA
• in some organisations a quasi-independent role

Senior Information Risk Owner
• formal Executive Director or Senior Management level role
• overall responsibility for an organisation's information risk policy
• accountable for information risk across the organisation

Lawyer/Legal
• **may be drafting a data sharing agreement or a license if you're making data available to others**

Information Governance Lead
• responsible for the use of information across the organisation
• may occasionally focus on personal data
• will also have an interest in the sharing of non-personal data

COMMUNICATIONS

Communications Officer
• could help shape project business case and internal communications

FINANCE

Finance officers
• Will know about the functions of all the organisation and will likely have reporting data and knowledge of what everyone does
• Internal audit team could be particularly helpful

PAS 183 Smart cities. Guide to establishing a decision-making framework for sharing data and information services **- more detailed guidance on the roles within a broader smart city plan.**

Stakeholder Analysis / Power x Influence Grid - allows you to assess stakeholders on the basis of their project interest and influence, which in turn recommends how to manage them.

RACI Matrix - helps identify roles individuals play at different points of the process. Makes clear who is responsible or accountable for a decision as well as who should be consulted and informed about the decision.

DISCOVER

UNDERSTAND YOUR DATA SHARING ECOSYSTEM

As well as getting to know who can help you internally, you will want to understand your broader data sharing ecosystem beyond your organisation. This will include the potential external organisations who you could share data with, as well as other organisations which might be able to help support you, either through providing technical services to help you share or analytical services on the basis of data once it's shared.

Here we have elaborated the different types of organisations which have been highlighted in our research, however there may well be others which we haven't included.

DATA INFRASTRUCTURE

City Data Platforms
In some cities there are technology platforms which have been built to support the sharing of data between different organisations. You may find these useful in providing technology to share data or to find data from other organisations.
Examples - MK Data Hub, Fiware, CityVerve Manchester, Data Observatories

Data Catalogues
Your organisation may be using a data catalogue to publish open data.
Examples - CKAN, Socrata, Data Mill North

Data Marketplaces
Data marketplaces make available data from different organisations and data producers.
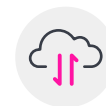**Examples** - ESD Toolkit

SERVICE SUPPLIERS

IoT supplier
IoT suppliers may be involved in collecting data on your behalf and therefore they may need to share data with you. Some may also offer additional services in terms of analysis and visualisation.
Examples - Cleanspace, Smart Parking, Siemens, IoTUK

Cloud Service Providers
Cloud Service Providers are used to host data, especially large datasets. They also allow the provision of other services such as data analysis and visualisation.
Examples - Amazon Web Services (AWS), Microsoft Azure

Data Producers
There are some organisations where data is a core part of their product offering. They usually make money through selling high quality data such as canonical lists or maps. These are both public and private sector organisations.
Examples - Ordnance Survey, Land Registry, Experian

Large Tech Providers
There are a number of large tech providers your organisation may have a relationship with and may be providing some services already e.g. database solutions.
Examples - IBM, Cisco, BT, Northgate, Siemens

Support Organisations
A number of organisations provide resources, tools and opportunities to meet those who are also undertaking city data sharing projects.
Examples - Future Cities Catapult, Digital Catapult, Nesta, the Open Data Institute, Universities

Data Science Providers
There is a growing number of consultancies who will provide data science services.
Examples - ASI, Universities

SUPPORT

Open Source Code
Open source repositories which may provide code to help in sharing the data you identify. This includes code for APIs etc.
Examples - Tombolo Digital Connector, GitHub, Libriaries.io

Data Standards
Data standards exist which may help the process of sharing as data can be more easily interpreted by different organisations.
Examples - British Standards Institute, City Data Standards Network, European Standards e.g Datex, Inspire, ISO

Forums and Communities
There are a number of forums and communities which you can access to find others who may be able to give support and advice on data sharing.
Examples - LocalGovDigital Slack, Twitter #finddata, Knowledge Hub Conferences

? Some questions to ask at this stage:
• Do you have a clear idea of the external organisations that could help you in your data sharing work?
• Do you know who externally could use your data?

Transport for London have a complex data ecosystem which they have built on the basis of a heritage of releasing open data. Their ecosystem includes organisations which provider travel planning services for people using transport in London (including apps like CityMapper), as well as those who incorporate travel data into their other services (e.g. Google Maps, Apple Maps). There is a range of organisations which re-use and repackage the data made available. **The Open Data Institute** have begun to **map TfL's eco-system**, as well as others, to show how open data is used to deliver public services and have also developed a workshop methodology too

Mapping Data Ecosystems (Open Data Institute) **- The Open Data Institute have put together a workshop methodology that will help you identify your broader data ecosystem.**

Service Innovation Handbook (Lucy Kimbell) - a resource with a number of tools designed to help people bring service design and innovation techniques into their organisation. In particular the "Mapping Innovation Ecosystem" toolkit provides a detailed workshop format to help you map your innovation ecosystem.

Platform Design Toolkit - Platform Design Toolkit is based on the tradition of Service Design Thinking and was born to overcome the limitations in traditional, linear, Business Modelling approaches. This methodology helps you model multi-sided, ecosystem based, platform models that are transforming the world of business and services.

IDENTIFY SPECIFIC DATA & SYSTEMS

Having understood the aim of your data sharing project and who is involved, you will want to identify the specific datasets and systems that will need to be shared. At first these might not be obvious and you may need to bring together those involved to help identify them.

This is likely to be an iterative process where conversations with colleagues and key contacts in other organisations may lead to the identification of further relevant datasets.

For those who are providing data, this step is important to make sure you're aware of all of the data which could be made available.

For those who are receiving data this step is important to show that you have checked you don't currently have the data you need and to identify the data you do have which you can use with the shared data.

The following approaches are some suggestions for how you could go about identifying the relevant internal data for your project - some of these we identified in our co-creation workshop. You might use a combination of these to identify the data and systems which are relevant to your project.

USE THE DATA YOUR ORGANISATION ALREADY MAKES AVAILABLE

You may be able to identify the relevant datasets you are looking for in a number of ways. You may have to produce statutory reports on services and these may point to underlying datasets held by your organisation. In addition, your organisation may also have a proactive agenda for publishing open data which will also highlight further datasets.

In terms of data which may not be made public, organisations will already likely have a process for data audit and cataloguing to be able to respond to the new responsibilities under the GDPR. For public sector organisations they also have a responsibility for making available a Information Asset List under the Re-use of Public Sector Information Regulations (2015). These lists are a helpful start to identifying relevant data sets.

IDENTIFY THE DATA THAT YOUR ORGANISATION IS LIKELY TO HAVE AND SEE IF YOU CAN LOCATE IT

If your organisation doesn't have a complete or up-to-data Information Asset List you may be able to identify data on the basis of the types of data it should hold. If you work in a local authority, the Local Government Association has put together a list of functions of local government and the associated data and standards for these functions.

IDENTIFY THE SPECIFIC DATA WHICH IS USED IN DELIVERING THE SERVICE

Engagement with the service team and the analysts which support them will help reveal the data that they use and have available. This engagement could take the form of individual discussions or workshops.

In a recent project with Essex public sector organisations, Nesta has used a Customer Journey Mapping methodology to map the interactions in the service and therefore the data and systems used at each stage.

IDENTIFY THE DATA THROUGH THE SYSTEMS YOUR ORGANISATION USES

Engagement with the team responsible for the systems in which your data is held (in many cases the IT team) may help identify relevant data. They should be able to identify the systems used within the organisation, who uses them and for what purpose. You could also identify this information on the basis of a survey of the systems used by a particular team or by observing their work.

The IT team may be able to provide you with documentation about the schema of each of the databases and help you create a more technical 'data dictionary'.

TRY AND DESCRIBE THE DATA THAT YOU NEED AND SEE IF ANYONE RECOGNISES IT

Another approach you could take is to describe the type of data that you expect to find or that would be useful to you. Having detailed this you could circulate this around your organisation in an appropriate way perhaps through an analyst or IG network to see if anyone is aware of this information. Alternatively you could reach out directly to the department who you think is most likely to hold the data.

CAPTURE DATA ABOUT DATA

Data about data is known as 'meta-data' and at this stage you will need to start collecting meta-data about the data of interest to you in your data sharing project. In some cases, the collective 'meta-data' is known as a 'data dictionary'.

The following is a (non-exhaustive) list of data that you may find helpful to capture about your data. On the particular topic you're interested in there may be a recognised standard for the meta-data that will help you organise the information.

- **Title of dataset** - The formal title of the data set
- **Description** - what topic does it cover etc.?
- **Personal data** - Does the data include any personal data as covered by the DPA?
- **Owner** - who in the organisation owns the dataset?
- **Uses and users** - how is the data currently used and by whom in the organisation?
- **Location** - Where is the data kept? Is it possible to capture a URL of the location and link to it directly?

- **System** - What system is the data held in?
- **Format** - What format is the data saved in? Common formats include .csv, .xlsx, .doc, .json, .pdf
- **Time period** - What time period does the data refer to?
- **Frequency of update** - How frequently is the data updated?
- **Publication status** - Is the data published in any form and if so where?

Having started to create a list of relevant internal data you may want to liaise with the person responsible in your organisation for the Information Asset List to see how the list you have curated can be incorporated in a broader list of information held by the organisation.

In **response to the Shakespeare Review of Public Sector Information**, the UK Government committed in 2013 to create a **National Information Infrastructure** (NII). The NII was described as containing "the data held by government which is likely to have the broadest and most significant economic and social impact if made available and accessible outside of government". In order to create this list, the Cabinet Office team charged with developing this led a collaborative process to create an inventory of data held by government. This list was first published on Data.gov.uk in 2013 and updated in 2015.

Using Customer Journey Mapping as a Tool for Data Discovery (Nesta) - You may want to use a tool from Service Design to better identify the specific datasets and systems

Data dictionary tools - A database of technical tools to support the development of data dictionaries.

DISCOVER

UNDERSTAND WHAT DATA OTHERS HAVE

As well as knowing what data your organisation has you will also need to know what data others have. You may already have a good idea of this, but if not here are some approaches that might be helpful.

If you are providing data, an organisation may have already reached out to you and it might be relatively straightforward to understand the data they already have and what they are looking to get from you.

If you are looking to receive data you may already have a clear idea of which organisation may be able to provide data. Alternatively you may need to identify specific datasets and organisations.

FIND AN EXPERT

In reality knowledge about data that could be available is built up by people in the course of their career. You may have a knowledgeable expert about data in your organisation - perhaps one of your analytical community has a good knowledge of the data which is held by other organisations? Perhaps there is someone who has been involved in the field who is more than willing to chat about the data that's available. It's always worth talking to more than one person as they will have a different perspective.

INTERNET SEARCH

Unfortunately there is no current catalogue of data held by public and private organisations which is openly searchable. However some open data is discoverable online - Google has recently launched a search engine for open data - toolbox.google.com/datasetsearch. This open data search might highlight organisations who hold data. Academic organisations and central government are key data holders and so you may want to search their websites specifically.

A more focused approach could be to search on the websites of the organisations you have identified as part of your data ecosystem. If they are a public body they may have a specific part of their website for data. If there have been Freedom of Information requests for the data previously these may be published on their website as well.

APPROACH DATA MARKETPLACES, CITY PLATFORMS AND VENDORS

Data may be available through data marketplaces or through a city data platform. These intermediaries are a developing industry and so the amount and availability of data they hold will vary.

NETWORK WITH THE COMMUNITY

As well as identifying an expert in the field there may be others in the wider community who are able to help. This could mean networking in person at a conference or meetup or online in a relevant forum that you've identified through your ecosystem analysis.

Local multi-agency networks and meetings might also be a useful opportunity to identify data. These types of meetings might include Local Enterprise Partnerships or where organisations in your Combined Authority meeting to discuss issues.

There have been a number of attempts to create canonical lists of the data held by the public sector at both the national and local level. These have been taken forward by various groups and organisations. Specifically in relation to Smart Cities there are a number of underlying models of the data which are being developed:

- BSI through it's PAS 183 have made available an underlying concept model
- FIWARE or FI-WARE is a middleware platform, driven by the European Union, for the development and global deployment of applications for Future Internet
- The Local Government Association has produced a Local Government Services list
- The Government Digital Service also undertook a programme of developing Registers of government data. A register is a list of information that is managed and approved by government. For example, the country register lists all countries recognised by the Foreign and Commonwealth Office.

Finding Open Data (the Open Data Institute) - the Open Data Institute has curated a list of resources where you might be able to find open data published by UK institutions.

23

DATA SHARING CANVAS

This template could be used in an inception workshop which you could run at the beginning of your project, or in a workshop at the end as a summary of your discovery phase.

This template is based on the Open Data Challenge Series template by Nesta and the Open Data Institute.

INTERNAL STAKEHOLDERS

Summarise the output here from **'Identify internal team'**

EXTERNAL ECOSYSTEM

Summarise the output here from **'Understand your data sharing ecosystem'**

Organisation Aims

User Needs

Summarise the output here from **'Clarify your reason for sharing'**

Summarise the output here from **'Identify specific data and systems'**

Summarise the output here from **'Understand what data others have'**

INTERNAL DATA

EXTERNAL DATA

A CITY DATA SHARING PROCESS

PREPARE
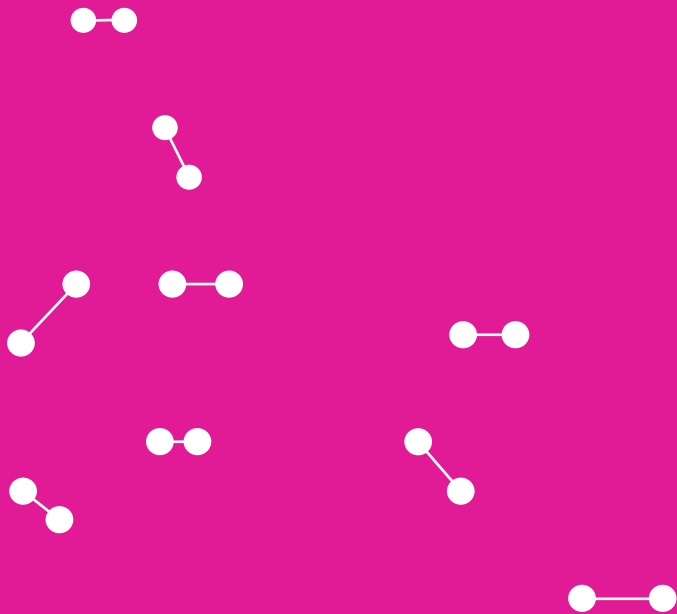
PREPARE
BUILD YOUR BUSINESS CASE

Your organisation is likely to require a formal business case for sharing certain types of data. This is the key opportunity to get clearance for your project from senior members of the organisation.

Different organisations will have different approaches to the development and presentation of a business case and there will likely be formal documents and processes which you will have to complete. Given this, it's best to establish what the expectations are for the development of your business case early and who specifically needs to be consulted. Perhaps you could undertake a prototype of the business case early in the process to identify areas where you will need to do more investigation.

HM Treasury guidance for the development of business cases in Government uses a 5-case model. This type of approach is usually used to develop business cases for spending on large capital projects such as infrastructure so not all elements are relevant for a city scale data sharing arrangement. In the following table, we have given some initial starting points for consideration for incorporation in a business case using this model, both for where you will be providing data and where you will be receiving it.

5 CASE MODEL

Considerations when you are providing data

Considerations when you are receiving data

1 **Strategic Case** - shows that the proposed data share fits with your organisations strategy, sets out why it's needed and what the expected outcomes are.

- A lot of this thinking will have been developed in the Discovery Stage of the project. However you may want to update this on the basis of the additional information you have been able to find.

2 **Economic Case** - demonstrate that the data share optimises public value

- Can you quantify the overall costs and benefits?
- In terms of costs, what is the cost of technical changes needed in order to provide the data? Are there any additional legal costs in terms of developing licenses?
- In terms of benefits, can any of these be quantified and over what timescale?

- Can you quantify the overall costs and benefits?
- In terms of costs, what is the cost of technical changes needed in order to receive the data? Does any further cleaning of the data need to be undertaken and how much will this cost?
- In terms of benefits, can any of these be quantified and over what timescale?

3 **Commercial Case** - this sets out any procurement strategy and the contractual arrangements needed

- What license will the data be made available under and are there any other legal considerations?
- If there is a charge associated with the data share and have you identified what the charging structure would be?

- Have you appraised all the possible different sources of similar data and what are their relevant benefits and disbenefits?
- What are the legal considerations in terms of the licenses you will be asked to take on?
- Is there a cost associated with the data?

4 **Financial Case** - demonstrates that the data share is affordable

- Can you detail the precise costs of the data share for your organisation?
- Who is going to pay for the data share? Do you need to find further funding for this or will it pay for itself?

5 **Management Case** - demonstrates that the data share is capable of being delivered successfully, in accordance with recognised best practice

- What are the project management arrangements for the data share? What methodology will be used?
- What are the arrangements for the ongoing monitoring of the data share?
- Are there any contingency plans for risk management?

Green Book and Business Case Guidance (HMT) - If you are working within a public sector organisation you may already have processes for developing a business case. However HM Treasury produce guidance on building business cases for spending public money which you might find helpful.

Business Model Canvas - This is a one page tool which may help to structure your thinking about the business model which underpins the data share, in particular if you are providing data as a product.

PREPARE

UPDATE LEGAL, PRIVACY AND ETHICAL APPROVALS

Part of the preparation process for sharing will also be to make sure that you have the appropriate legal powers to provide or receive the data. Even though the data we have been considering in this toolkit is non-personal it is good practice to ensure you go through the Privacy Impact Assessment process. Increasingly, organisations are encouraged to consider the ethical implications of the data they use - this holds true for data sharing as well.

Practice in this area is continuing to develop and so the information presented here should be seen as a starting point and you should get tailored advice from internal experts such as your legal department and information governance function.

LEGAL POWERS FOR SHARING

There are likely to be a number of legal considerations in sharing data and you will want to get legal advice early on in the process to understand what these are. There are likely to be specific considerations dependent on the type of organisations involved in the data share, the type of data being shared and the value exchange which has been identified (especially if there is payment involved). These legal considerations are separate but related to ensuring that you're not sharing personal data (covered subsequently).

There are overarching legal considerations which will shape your approach to sharing data. For instance you will want to check you have the legal powers for sharing the data identified and that you're keeping within your general responsibilities for making data available if you are a public sector organisation. Working with your legal department you will want to make sure that you have either express or implied powers to share the data you've identified. For example, there may be a specific relevant legal gateway for sharing the data or sharing the data may be covered in the context of you delivering your statutory functions.

Public sector organisations should be aware of their obligations under the Re-use of Public Sector Information Regulations (2015). This relates to the data which they produce as part of their 'public task' but which are not already covered by the Data Protection Act (2018) and General Data Protection Regulation (2016). There are certain obligations to the provision of data for re-use, the charging regimes which they are able to use and recommended licenses.

On the understanding that your organisation has already collected this data you should have already established your legal powers for collection. There may be particular legal considerations if you are not the organisation which is involved in the collection of data. For instance, you may need to ensure that you have ownership of the data which is collected on your behalf.

As the data will be new to your organisation you may need to go through new legal, privacy and ethical approvals processes.

You might come across the following legal documents used in your organisation:

• **Memorandum of Understanding (MOU)** sets out a common understanding of a situation or process between two or more organisations in a formal manner but is likely not to be legally enforceable.

• **Data Sharing Agreement / Information Sharing Protocol** - This document summarises the overall framework for the secure and confidential obtaining, holding, recording, storing and sharing of information between participating partner agencies or organisations

• **Service Level Agreement (SLA)** - A service-level agreement is a commitment between a service provider and a client.

• **License** - This is an official document which allows you to use a dataset. It does not necessarily transfer ownership of the data and will probably specify the circumstances in which you can used the data, including any related charges.

The **Information Sharing Gateway (ISG)** - The Information Sharing Gateway (ISG) has been developed by a sub-group of organisations in the Lancashire & Cumbria IG Group in order to improve and modernise the administration and risk assessment of information sharing in the public sector. It allows organisations to keep all of their Information Sharing Agreements in an online repository. It incorporates a number of workflows which allows organisations to create and manage agreements and all sharing partner organisations have visibility of sharing agreements as they are being created and can contribute them.

Re-use of Public Sector Information Regulations (2015) - This guide to RPSI includes guidance on requests for re-use of information, charges and licenses which should be applied.

Data Sharing Code of Practice (ICO) - This code has not been updated since the Data Protection Act 2018 became law. The ICO are working on updating the code and have launched a call for views. The updated code will explain and advise on changes to data protection legislation where these changes are relevant to data sharing. It will address key aspects of the new legislation including transparency, lawful bases for processing, the new accountability principle and the requirement to record processing activities.

PRIVACY

Although you're likely to have already established that the data you're wanting to share is non-personal, it is good practice to ensure that you undertake a Data Protection Impact Assessment. Depending on your organisation this might also be standard practice and part of the information governance policies.

You might come across the following documents used in your organisation in realation to privacy:

- **Privacy Notice** - This is a statutory notice which has to be published so that individuals know how an organisation is processing their personal data.

- **Data Protection Impact Assessment (DPIA)** - a process to help you identify and minimise the data protection risks of a project. The ICO sets out a checklist to help you understand whether you need to undertake one of these and what steps to follow.

As part of **CityVerve**, Manchester's smart city demonstrator, the team set up a process for undertaking Privacy Impact Assessments. They used a Privacy Impact Assessment screener to engage colleagues to understand when they need to complete a document and provided support and guidance in completing the detail. Due to the complexity of the consortium, the team took legal advice to identify who was best placed to sign off the completed PIAs. A complete PIA assessment was needed in every instance of data sharing, regardless of whether the data was explicitly personal or not. A lesson from the project was to start the PIA process at near the beginning of the project rather than to start later on.

Data protection impact assessments (ICO) - DPIA (also know as a Privacy Impact Assessment) is a process to help you identify and minimise the data protection risks of a project. The ICO publishes best practice guidance which includes useful checklists and approaches.

ETHICS

There is increasing interest in the ethics of data use, again especially in relation to personal data. A number of guides and tools have been developed in particular by DCMS.

Data Ethics Framework (DCMS) - Public sector organisations should use the Data Ethics Framework to guide the appropriate use of data to inform policy and service design.

The Ethics of Data Sharing: A guide to best practices and governance (Accenture) - Document setting out some of the specific ethical considerations of sharing data including a useful approach to risk management.

Data Ethics Canvas (Open Data Institute) - The Data Ethics Canvas is designed to help identify potential ethical issues associated with a data project or activity. It promotes understanding and debate around the foundation, intention and potential impact of any piece of work, and helps identify the steps needed to act ethically.

PREPARE
PREPARE DATA

Having identified the specific datasets which you are wanting to share you will want to make technical preparations for sharing.

If you are providing data you will want to make sure that you are aware of all the needs of those who will be using the data.

If you are receiving data you will want to understand more detail of the data you are receiving and how to prepare your existing data so that it can be useable with the new data.

CLARIFY YOUR USERS' NEEDS

It's best to understand what the users of your data need in order to be able to develop the appropriate specification for the data.

This is where it's important to have developed working-level relationships between those organisations involved in the data share. Perhaps you could collaborate on a specification for the data or find a data standard which is appropriate.

Perhaps you can talk to the current users of the data to help build content for the documentation. Through this process it would be important to document those aspects of the use of the data which are taken for granted.

MAKING LINKS

One way of increasing the usefulness of the data being shared is to include the most relevant shared identifiers so that it can be easily linked to other datasets in the receiving organisation. Some common examples of shared identifiers include:

- **Address data** - In particular where you have data on properties you may want to use the Unique Property Reference Number and / or the Postcode depending on need

- **Administrative and Statistical Geographies** - Administrative geographies are defined by government and usually reflect the organisational make up of local government. The Office for National Statistics produces statistical Georgraphies. Both of these are available through a Gov.uk Register
- **Companies** - A common unique reference indicators that is used to identify companies is the D-U-N-S Number from Dunn and Bradstreet.

You may or may not need to consistently use a data standard or shared identifiers dependent on whether the organisation which is using the data has data matching tools. These tools can match datasets in the absence of standards or identifiers and are becoming increasingly widely used.

You will need to ensure that you have the right license to incorporate data which you have not created yourself into a dataset which you subsequently share with another organisation.

QUALITY ASSURANCE

You are likely to have internal processes already to ensure data quality and should spend time understanding how these can support the process of preparing the data for sharing.

Use the data yourself
If there are currently no users for the data internally, or it hasn't been used so far, perhaps you could undertake some exploratory data analysis on the data so as to understand it's limitations and foibles. Visualisation of the data may be a quick way to identify any anomalies with the data.

Clean the data
If there are inadequacies highlighted in the data you will need to make plans and potentially seek resource to make amends. Some of these may be relatively straightforward (e.g. using a consistent naming of entities throughout the data and deduplication) however some may require changes to data collection in the longer term.

Be transparent about data quality and remedies
You can summarise the quality of the data in a Data Quality Statement. You may want to specify the process by which there is continual evaluation of the data to make sure that it is fit for purpose.

EXPLAIN THE DATA

As well as the actual data set you will want to develop some supporting documentation that explains the content of the data and the technical use of the data.

Involve current experts

Domain knowledge is essential here to understand the complexity, meaning, history and future of the dataset you're trying to share. Make sure that you get rid of jargon and acronyms so that the dataset and supporting documentation is as understandable as possible to other users.

There are several key aspects to explain in your documentation about the data and some of this may already be incorporated into the meta data:
• How the data was captured and why (Provenance)?
• Explain the processes of data verification and validation that are undertaken in the collection and current use of the data
• How confident you are in the accuracy and completeness of the data?
• How up-to-date the data is (Currency)?
• Frequency of update of the data.
• Data standards that are used in the data
• Whether there is any embedded IP any of the data which re-users should be aware of.
• Format of the data.

ESD Toolkit / LGInformPlus - Produced by the Local Government Association this toolkit gives provides details of many different data standards used in local government.

Schema.org - is a collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond. You may be able to find a standard for your dataset here.

Tim Berners-Lee 5 Stars of Open Data - Sets out recommendations for publishing machine readable open data.

PREPARE

UPDATE YOUR TECHNICAL ARCHITECTURE

You will need to spend some time designing and testing the technical architecture for sharing the data. This means both the systems used to share data but also the related data format in which the data is made available.

The architecture you choose will depend on:

1. The capabilities of your internal systems
2. The capabilities of the systems of the other organisations you are sharing data with
3. The nature of the data being shared

Your IT colleagues will be able to advise about how your current technical architecture can be set up to enable you to be able to share the data you have identified. Below is a quick summary of some of the issues you will need to consider at this stage, however you should seek more technical advice.

CONSIDER CURRENT SYSTEMS AND YOUR BROADERS STRATEGY

Can your current systems be used to undertake data sharing and if so how much does that cost in comparison to the alternatives? Does tailoring your current system have implications in terms of capacity and training of those who will use it internally?

The precise nature of the technical aspects of the data share may mean that there will need to be further investment in your systems and this cost will need to be incorporated into your business case.

SUSTAINABILITY

A big consideration in terms of the technical architecture approach you take is whether you are sharing the data on a one-off basis or ongoing. How sustainable are your plans for the technical architecture of the data share? How will the ongoing maintenance of the system be financed?

IN-HOUSE OR OUTSOURCED

Further considerations are whether you should develop the data sharing technical architecture in house, get a supplier to develop a bespoke system or buy an "off the shelf" solution; the implications of this in terms of maintenance and up-keep; whether you have the skills internally to undertake any development; whether it's better to train employees or use an outside supplier; and how this decision fits with your broader Digital and IT strategies for the organisation.

TECHNICAL SECURITY

You will also need to establish whether your proposed approach meet your organisational technical security requirements, and how it fits with national guidance.

New regulations on the Security of Network and Information Systems were passed into law in 2018 and a number of organisations are being set up to provide oversight and enforcement of these.

As part of the Tombolo project, Future Cities Catapult developed a tool to make it easier to share and combine different sources of data. The **Tombolo Digital Connector** is an open source software library that enables users to seamlessly combine different sources of data in an efficient, transparent and reproducible way.

At present, the tool can be used to combine information relating to mobility, the environment, air quality and even the street network. Because it is open source, the Tombolo Digital Connector can be extended by anyone to meet their needs for data sharing and combination. It is designed to work on a range of operating systems including, MacOS, Windows and Ubuntu.

In our research for this toolkit we identified a number of patterns of how city data is currently being shared. Arguably the most benefit from data comes when using an approach which enables less human effort in the sharing of data itself. This then means that people can be redeployed to analysis and implementation.

1. SHARING DATA IN PERSON

Perhaps one of the most frequent ways in which city data is shared is in person. This could be at a multi-agency meeting or bilaterally.

2. ACCESS TO TWO UNCONNECTED SYSTEMS

Data from one system is transferred to another by an individual who has access to both systems. However, the two systems don't connect technically. This can be a time consuming process.

3. ONE-OFF SHARING OVER THE INTERNET

Data can be shared with specific individuals and organisations through an email attachment, on an FTP server etc.
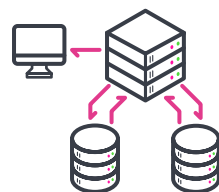
4. SHARING USING A THIRD PARTY

Data between a number of organisations who lodge it with a third party. This has benefits in that the third party can offer additional functionality as well as analytical services. This is a pattern of sharing which can also happen internally.

5.SHARING DIGITAL DATA IN ANALOGUE WAY

In a number of instances in the public sector data is still shared by printing out information system and then sending to another organisation / individual for input on their system.

6. ONE-OFF SHARING THROUGH PHYSICAL MEDIA

Large data sets are still often shared using media such as USB stick or DVDs. These are sometimes sent through the post.

7. OPEN PUBLICATION OF DATA

Data is published on the organisation's website in an open format and with an open license. Local Authorities have responsibilities to publish transparency data openly.

8. SHARING USING AN API

Data is made available through an Application Programming Interface which allows machine to machine transfer of data on an automatic basis. Access to the data can be controlled through a registration process.

Technology Code of Practice (Government Digital Service) - The Technology Code of Practice is a set of criteria to help government design, build and buy better technology. It's used as a cross-government agreed standard in the spend control process.

API technical and data standards (Government Digital Service) - GDS have put together detailed guidance on the design of APIs for government covering aspects such as datafields and documentation.

Guidance (National Cyber Security Centre) - NCSC's guidance is aimed at helping UK government departments, agencies, the critical national infrastructure and its supply chains protect their information and systems. It also has relevance for local government and the wider public sector. They have recently published a pattern for the safe import of external data.

NIS Directive and NIS Regulations 2018 (HMG) - Further information about the Security of Network and Information Systems Regulations.
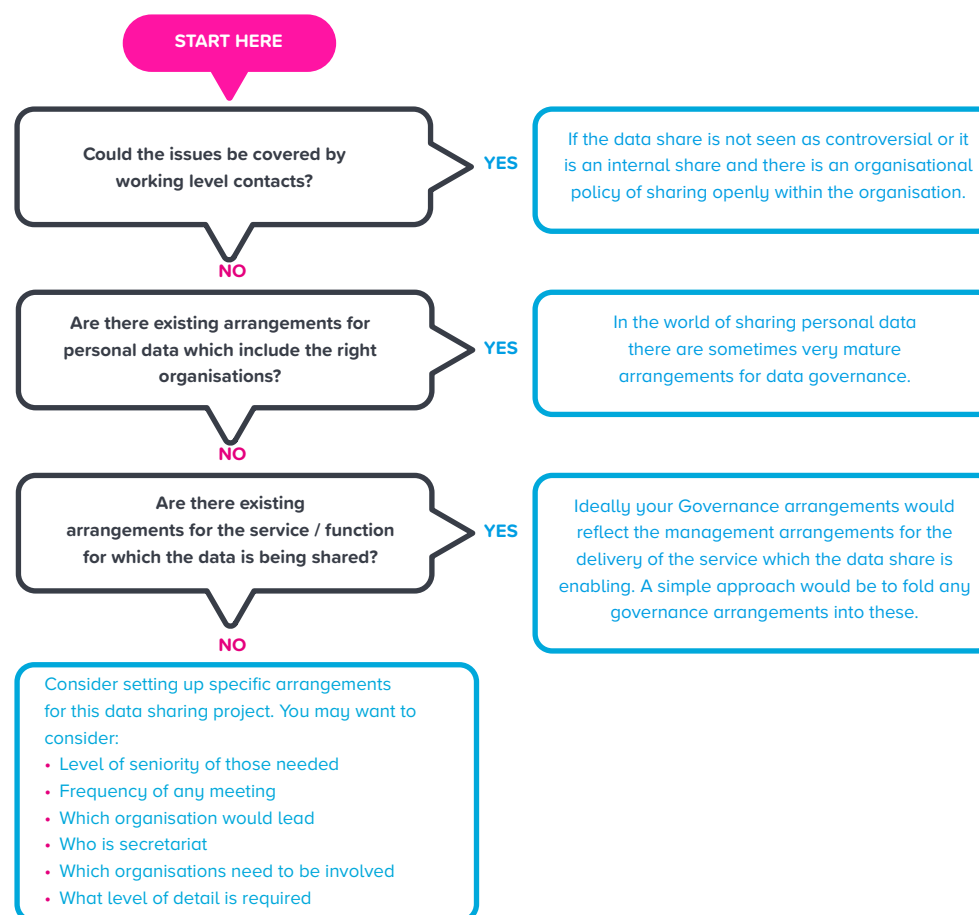
PREPARE
CONFIRM GOVERNANCE

Governance arrangements for data sharing are important where there is ongoing sharing of data. However, this is a developing area and arrangements can vary from organisation to organisation and dependent on the types of data which are being shared.

By Governance we mean all of the arrangements for ensuring the oversight and accountability of the data sharing arrangements. This includes the processes for programme and project management, including formal meetings with senior members of the organisations sharing data and identifying signatories for formal contracts.

Formal governance arrangements are common when sharing personal data on an ongoing basis. In these circumstances, some of these governance arrangements are set out in the UK Data Protection Action 2018, such as the role of the Data Protection Officer.

We have put together this flow chart to help you ask questions about the type of Governance arrangements that you need and to design the right option for your organisational circumstances. This is a work in progress and we would value feedback.

If your organisation has an Information Governance function they will be able to help you design the appropriate Governance arrangements for your data sharing programme.

START HERE

Could the issues be covered by working level contacts?

YES — If the data share is not seen as controversial or it is an internal share and there is an organisational policy of sharing openly within the organisation.

NO

Are there existing arrangements for personal data which include the right organisations?

YES — In the world of sharing personal data there are sometimes very mature arrangements for data governance.

NO

Are there existing arrangements for the service / function for which the data is being shared?

YES — Ideally your Governance arrangements would reflect the management arrangements for the delivery of the service which the data share is enabling. A simple approach would be to fold any governance arrangements into these.

NO

Consider setting up specific arrangements for this data sharing project. You may want to consider:
• Level of seniority of those needed
• Frequency of any meeting
• Which organisation would lead
• Who is secretariat
• Which organisations need to be involved
• What level of detail is required

The Government has recently announced the development of a **Geospatial Commission** to oversee to maximise the value of all UK government data linked to location. This is a very high-level governance arrangement for data held by a number of government organisations which create and hold location data. The organisations involved include HM Land Registry, the Ordnance Survey, the British Geological Survey, the Valuation Office Agency, the UK Hydrographic Office and the Coal Authority. The Commission is in the process of being set up but will include two joint chairs and a number of independent commissioners. The first task of the commission will be to establish how to open up freely the OS MasterMap data to UK-based small businesses.

A CITY DATA SHARING PROCESS

SHARE
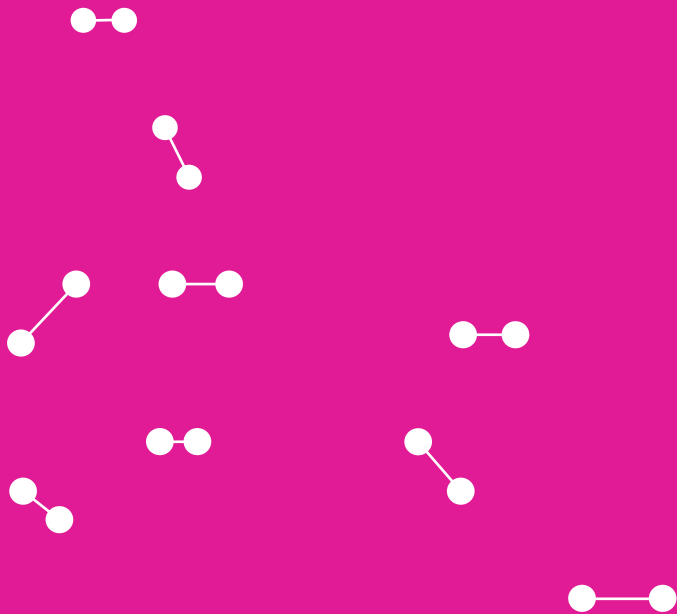
SHARE

SUPPORT A SHARING CULTURE

Having begun to share data there are a range of things that you may want to do in order to support the continued sharing of data and to build a broader culture of collaboration within your organisation and beyond.

There are lots of different opinions and theories about how to change organisational culture and individual behaviour. Clearly it's also not your place to use a data sharing project to try to change the entire culture of the organisations who are involved. However, we've put together some thoughts from our research about how to support a broader cultural collaboration so as to make data sharing easier:

TRANSPARENCY

• Can you make available any of the documents you have created in the process of setting up the data sharing?
• a blog on your intranet could help raise awareness of the project or perhaps you could blog on your org website or sites like medium. com.
• can any parts of your code be published on Github?

TRAINING

• Are there ways in which you can incorporate key aspects into other training docs (Mandatory IG training in your organisation)?

TECHNOLOGY

• What technology can you put in place to help the two teams stay in touch? Email may be the simplest way but other options such as Slack could build a closer relationship.

MEETINGS AND EVENTS

• Can you talk about the process of developing the data share at a 'Show and Tell' or some other group event?
• Are there any meet-ups where you can talk about your work?

Culture mapping (Dave Gray) - This tool helps you identify the aspects of your current culture and can also be used to map a desired future state.

Open Leadership 101 (Mozilla) - This course describes the behaviours and culture that Mozilla uses which it calls "working in the open". Many of these approaches may be helpful in supporting a broader sharing culture in your organisation.

Making things open, making things better (Mike Bracken) - Blog setting out the rationale for an early GDS to use open source and more open approaches.

NEXT STEPS
FUTURE CITIES CATAPULT SUPPORT

This toolkit sets out a high-level process for designing and running a city data sharing project. The activities are likely to take some time and it is not always a straightforward process.

Having been involved in a number of data sharing projects, the Future Cities Catapult team has experience of working with organisations at all stages.

Future Cities Catapult's Data Science Team can help you and your organisation scope out the feasibility of data sharing project and support you through every aspect of the process.

We are available to run initial inception workshops, help you identify relevant datasets which might be of use, point you in the direction of best practice and advise on some of the technical solutions you might want to consider. We can also work with you to help build a broader culture of collaboration and sharing within and across organisations.

If you would like to talk about the possibility of working with us on any of the above, or if you have a specific project you're thinking of doing, please get in touch with Yalena Coleman, Data and Demonstrators Account Manager on ycoleman@futurecities.catapult.org.uk.

ACKNOWLEDGEMENTS

FUTURE CITIES CATAPULT TEAM

Ed Parkes
Content Lead

Tom Leaver
Project Manager

Felipe Caro
Design and Illustration

Lee Markham
Narrative Design

David Altabev
Director, Data and Demonstrators

Dr Joseph Bailey
Data Science Team Lead

INTERVIEWEES

As part of the research for this Toolkit we spoke to the following and would like to thank them for their contributions:

Adrian Slatcher
Manchester City Council

Alan Moore
thinkWhere

Andrew Beechener
Republic of Things

Conor Dumigan
Analytics Engines

Carl Rodrigues
Streamr

Geoff Snelson

Dr Jacqui Taylor
FlyingBinary Limited

John Gibson
ASI

Jovian Smalley
Information Commissioner's Office

Fran Bennett
Mastodon C

Hilary Simpson
Sleuth

Nick Allott
nquiring minds

Nicola Bryant
Digital Birmingham

Luke Smith
Technical Architect   University of Newcastle

Nicola Biggs
University Hospitals of Morecambe Bay

Paul Connell
ODI Leeds

Paul Hodgson
Greater London Authority (GLA)

Raj Mack
Digital Birmingham

Richard Vilton
Emu Analytics

Sam Roberts
Department for Digital, Culture, Media and Sport

Sarah Butler
Greenwich Council

Sarah Gates
Department for Digital, Culture, Media and Sport

Simon Navin
Ordnance Survey

Steve Skelton
Stockport City Council

Thom Townsend
Department for Digital, Culture, Media and Sport

Tom Symons
Nesta

WORKSHOP ATTENDEES

We held a workshop on 5th September 2018 to test an early version of the content of this Toolkit. The following attended this workshop:

Abbas Lokat
GeoPlace LLP

Adrian Slatcher
Manchester City Council

Camilla Bertoncin
Nesta

Daniel Taylor
Sheffield City Council

Elizabeth Baker
Manchester City Council

Giles Robinson
Sheffield City Council

Greg Yiangou
Hertfordshire County Council

Ingrid Koehler
Local Government Information Unit

Jacqui Taylor
FlyingBinary Limited

James Maddison
The Open Data Institute

Jennifer Sayer
Humanising Autonomy

Jose Maria De Uriate Diaz
CypherOrdin Inc.

Laurie Platt
Sheffield City Council

Maciej Ziarkowski
Local Data Company

Mareike Schmidt
Q-Bot

Mark Cridge
mySociety

Michael Mulquin
 IS Communications Ltd

Michelle Eaton
Nesta

Nicola McHugh
Sheffield City Council

Rhoswen Hoath
Ordnance Survey

Ryan McMaster
Metasphere Ltd

Stuart Cole
Oxfordshire County Council

Waverley Coquet
The Open Data Institute

There are a number of organisations which have useful programmes that relate to the issue of data sharing:

- City Standards Network
- British Standards Institute
- Information Commissioner's Office
- The Open Data Institute
- Local Government Association
- Nesta
- The Data Lab

FEATURE BACKLOG

This is the first version of this City Data Sharing Toolkit. During its development we have identified some further improvements that we'd like to incorporate in a next version.

- **Expansion / refinement of the data sharing process** - to include a review and retirement stage for data sharing arrangements

- **Assessment of tool quality/utility** - In a subsequent version it might be possible to flag the quality and usefulness of the tools which we have identified

- **Non-PDF** - We'd like to explore other formats for making the Toolkit available in any subsequent version

- **Training and workshops** - There may be an opportunity to develop further training or workshops to support the use of the Toolkit

- **Interactive tool** - A further version of this tool could include an online interactive version allowing some of the tools and activities to be completed online.

- **International editions** - In different languages / tailored to different cultural and legal contexts

CITY DATA SHARING TOOLKIT

About Future Cities Catapult

The Future Cities Catapult's mission is to help UK firms develop innovative products and services to meet the changing needs of cities, and to sell them to the world. In doing so, we will support the emerging Advanced Urban Services sector in becoming an enabler of national productivity and a central plank of the UK economy.

We develop ambitious and impactful new projects that leverage our established strengths as a neutral convenor, helping buyers better articulate their needs to the market, working with suppliers to respond to those needs, engaging academic leaders to translate the latest R&D into application, driving the dissemination of evidence that unlocks investment and adoption, and provoking creative disruption in services and systems overdue for innovative transformation.

In short, we are dismantling the barriers to market success, ensuring that British innovators and entrepreneurs are positioned to enjoy a disproportionately large share of the global market for advanced urban services – a market estimated to value $1 trillion globally by 2023.

If you have any comments about this document or if you'd like to get in touch with the authors please email **jbailey@futurecities.catapult.org.uk**.

Visit our website
www.futurecities.catapult.org.uk

Follow us on Twitter
@FutureCitiesCat

Or send us an email
info@futurecities.catapult.org.uk

CATAPULT Future Cities