Local Data Review:

Open Data Approaches and Practices

Version 1.1

4th December 2020

Jess Hambley – Lead Analyst, Digital Analytics

Chris Pope – Principal, Digital Analytics

GREATER MANCHESTER
DOING THINGS DIFFERENTLY

Executive Summary

Greater Manchester aims to be one of Europe's top digital city regions. A key part of this aim requires Greater Manchester to make the best use of its public sector data assets. Opening up data for re-use can aid productivity growth through commercialisation and innovation. It also acts as a signal for our ambitions to develop and grow our digital businesses and our digital talent.

The Local Data Review helps to realise these ambitions, as laid out in our Local Industrial Strategy and our Digital Blueprint. It's aim is to identify the challenges to making more data open. By doing so, we can identify mitigating actions that increase the usability and use – and thereby the value – of our major public data assets.

This document brings together key findings on the key challenges and potential mitigations around open data, as highlighted through a collection of 25 open data case studies, an open data workshop, and wider research. These findings are split into distinct themes across five different sections of the report.

Section 1 first summarises the historical and strategic context of open data in the public sector. It outlines how open data originally emerged in the public sector under the government transparency agenda. However, in more recent years, open data has begun to be understood as an asset in itself, with the potential to add value to the local and national economy. This has allowed open data to rise in prominence within a range of national and local government policy and strategy. As a result, this section of the report outlines the relevance of this piece of work and links it to wider current strategic priorities and objectives.

Building on this historical and strategic context, as well as on the findings of the case studies and wider research, Section 2 moves on to consider what we have learned about open public sector data more broadly. It outlines how many open data initiatives follow a general evolution from a focus on data **quantity** to data **quality** to data **usage**. At each of these stages, a business case is needed to drive an open data programme forward. These different stages tend to involve broadly similar features, which allows for a general consideration of where Greater Manchester is currently at in relation to these stages. With different public sector bodies across the region arguably sitting at different stages along this progression, this section suggests that creating a **roadmap** to take these organisations from stage 1 to stage 4 would be valuable.

Section 2 also considers the wider **value** of open data, particularly noting its potential as an **economic enabler** that creates benefits downstream from those who actively interact with the data. This can prove challenging from a business case perspective, but it is important to recognise the wider secondary benefits of open data. Many public datasets have value to the private sector and have the potential to **drive growth and investment**. This section therefore outlines a range of priority data areas with high commercial potential for Greater Manchester: transport; business rates; socio-economics and demographics; planning; movement and activity; COVID-19; asset location; and infrastructure and risk management.

Having built up the evidence base to support the wider value of opening public sector data for re-use in Sections 1 and 2, Section 3 then looks at the main **issues and challenges** experienced by organisations when trying to open up their data. While challenges can be unique to each individual organisation, common themes run throughout a number of case studies and our wider research, which ultimately creates barriers to the long-term sustainability of open data programmes. These challenges are explored within the themes of: long-term sustainability; leadership and culture; data and data management; skills; matching supply and demand; and communications and awareness.

By exploring the experience of other organisations in approaching the challenges outlined in Section 3, we have been able to identify a range of **potential mitigations**. These are outlined in Section 4 of the report, which highlights actions that organisations could take to more effectively open up data and increase its usage. These recommendations are structured under a **MoSCoW framework**, which outlines the recommended features and approaches that an organisation **must** have, **should** have, and **could** have in order to effectively open more data.

Lastly, Section 5 draws on the range of findings from Sections 1 to 4 to outline a set of **open data deliverables** that might be needed in Greater Manchester. Deliverables are prioritised using the same MoSCoW framework outlined in Section 4.

From our work with local and national organisations, we have identified a series of nine actions that are of the highest priority in ensuring Greater Manchester can make good quality data more widely available for re-use. These actions are:

1. **Build on the success of open data projects in Greater Manchester by prioritising and opening up more, simple datasets that users want and need.** This can help to grow demand, and stimulate the view of Greater Manchester as an open, digital place.
2. **Develop and agree upon a set of basic standards for the highest priority datasets.** This should ensure data is consistent and comparable across Greater Manchester, adding value to the datasets and supporting their wider re-use.
3. **Organise a programme of work to support more open data releases, and to effectively identify and prioritise new data releases.**
4. **Create an easily accessible place to share open data for each public sector organisation in Greater Manchester**, building on the good practice and learning of organisations like Trafford Data Lab, Salford City Council, Stockport Council, Wigan Council and others.
5. **Establish a basic central repository of metadata that describes all datasets held by the public sector in Greater Manchester.** This aim would be long-term and could be based on information asset registers. This action would help potential users to identify datasets that they might not have previously been aware of.
6. **Ensure open data is provided to at least 3 star level and in more than one format.** This means all data should be provided in a non-proprietary, open format that does not require any particular software package – for

example, using CSV format instead of Microsoft Excel's XLS or XLSX formats. This can enhance usability, and thereby usage, of new open data.

7. **Establish an inclusive governance framework for open data across Greater Manchester.** This is already in place at an organisational level across many Greater Manchester organisations, but there is currently no GM-wide governance framework specifically with an open data remit.
8. **Lay out an open data roadmap for the Greater Manchester public sector,** to guide organisations through the progression of simply opening data, to focusing on increasing data usage and value.
9. **Articulate clear, consistent messaging around the value and purpose of open data,** linking this to strategic aims from the Greater Manchester Digital Blueprint, Local Industrial Strategy, and Greater Manchester Strategy. This should be supported by a strong communication strategy to guide data releases.

It is worth noting that these deliverables are indicative at this stage. They are also largely based on an internal public sector perspective – what are the challenges we face in making data open? This is only one piece of the puzzle, however. It is vitally important that we also begin to understand the challenges data users face in accessing and using our data.

Our next steps will be to develop this understanding of the use of public sector data. Alongside this review of open data approaches and practices, we are running an external business engagement exercise. This work, conducted by Open Data Manchester, will seek to gain an understanding of the external perspective of public sector open data: the challenges in accessing and using open public sector; and the opportunities to make accessibility and use of our data easier. While this work will predominantly focus on the business and commercialisation uses of public sector data, the aim will also be to understand the innovative academic and third sector challenges and uses as well.

The list of priorities and actions in Section 5 will therefore be amended and added to, following this exercise. Ultimately, we will endeavour to develop a prioritised list of deliverables for Greater Manchester, that should help us to open more data; ensure that our data are useable and used; and create one of the best open data ecosystems in Europe.

Introduction

Background and context

Leading digital city-regions appreciate the value of their local data, using data to inform and improve public services, and releasing data for wider re-use. Greater Manchester aims to be one of the top five European digital city-regions, recognised globally for its digital innovation, as set out in the Greater Manchester Digital Blueprint. Data – including how it is collected, used, shared and valued – forms a key part of that ambition.

The Greater Manchester Local Industrial Strategy (LIS), released in June 2019, also supports these ambitions, with an aim to create an economy fit for the future. By opening up more local data for re-use, we can help to create and support prosperous communities, stimulate innovation and increase both productivity and earning power.

The importance of better quality and more value-rich data is central to this ambition. The value of data comes in its use, not in its storage. Furthermore, data is more likely to be used when it is usable – that is, when data is accessible, and of sufficient quality to use. As a result, releasing data for wider re-use is not simply about opening up information, but also about ensuring that information can be widely used.

With these ideas in mind, the LIS proposed the creation of a Local Data Review (LDR) to identify and address the barriers to opening up local public sector data for re-use. There are three workstreams in the LDR:

- **Analysis** – identifying some of the main challenges the public sector faces when opening up data for re-use;
- **Consultation** – identifying some of the main challenges external organisations face in accessing and using public sector data;
- **Review** – identifying deliverables, and actions to create these deliverables, to open more public sector data.

The analysis workstream explores the potential challenges and mitigations for opening public data from an internal, public sector perspective. The consultation workstream, by contrast, looks externally at the challenges faced by businesses in using open data, and explores the types of data that businesses need. The review part of the LDR will then bring these internal and external perspectives together, with the overarching aim of enhancing data usage and improving data usability across Greater Manchester.

This ‘Open Data Approaches and Practices’ review is part of the **analysis** workstream of the LDR. By exploring case studies of existing and previous open data projects from a local to international scale, it is possible to draw out the key challenges faced by organisations when opening data, as well as potential mitigation options. Bringing a collection of case studies together also identifies the types of widely used and requested data, which provides insight into the prioritisation of what data to make open within Greater Manchester.

Further to the case studies developed, the team also ran an open data workshop session on 10th September 2020. This session brought together partners from

Greater Manchester and around the UK, exploring the initial findings of the case studies. It also sought to add a degree of prioritisation of the work to be completed.

Methodology

This report summarises the variety of work carried out across the analysis stage of our Local Data Review, from initial literature review, to case study selection and analysis, to our open data workshop. Our approaches for each are briefly outlined below.

Literature Review

Initially, we conducted a literature review on opening up more public data through the IDOX-Knowledge Exchange System.

The literature review covered opportunities, risks and challenges of opening up more public sector data. There was a request for a particular focus on open data used to contribute to economic growth and productivity through commercialisation and innovation. The review included local, national and international resources to identify:

- Examples of commercialisation of public sector open data (either by the public sector itself, or by others), and use of open data in open innovation.
- Discussions of barriers to releasing data openly (legal, technical, skills etc);
- Private sector and others' use and experience of public sector data, including any barriers they regularly face;
- The value of public sector data, and in particular the value of releasing data;
- Evaluations of open data projects nationally or internationally, including around data quality; and,
- Existing and well-known open data sites, including different ways of sourcing open data.

This review provided us with some key background information in relation to the opportunities, risks and challenges of opening up public sector data to drive economic growth. It also provided some initial insight into the wider potential value of opening public sector data, and provided examples of where others have used open data to work in a new way.

Case Study Selection and Review

The literature review helped identify some initial case studies to review, such as Dublinked and Transport for London. Further case studies were then identified through a combination of discussions held locally, and desk research carried out through other sources and links. In total, 25 case studies have been carried out and reviewed by GMCA Research. Case studies have largely been selected based on their location and/or relevance to the objectives of the LIS and Digital Blueprint, which means that the findings are tailored towards Greater Manchester. However, the core findings are also widely applicable to open data projects elsewhere.

The team conducted desk-based research on our case studies to find out about the programmes and the teams behind them. There were several key areas that acted as the focus of the review: what the project, programme or team was, and how it was initially set up; the key challenges faced by those behind the programme; challenges with interacting with the data from a user perspective; and what lessons can be applied from these case studies to Greater Manchester. Where possible, a number of one-to-one meetings were arranged with relevant individuals and teams involved in our case studies. This provided further insight into the matters identified through our desk research.

The full list of case studies explored in this review are listed in Appendix 1, along with their relevance for inclusion in the study. Our individual case studies are available for download on our website here: https://www.greatermanchester-ca.gov.uk/what-we-do/research/research-digital/local-industrial-strategy-local-data-review/

Open Data Workshop

Upon completion of our case study review, GMCA Research ran an open data workshop. The workshop was split into two halves: the first session explored how different areas might work towards a better open data ecosystem in a general sense, while the second focussed on the actions Greater Manchester specifically needed to take.

Around 30 participants attended each session from a range of public sector organisations, both within Greater Manchester and across the wider country. Many of these participants were leaders in the field of open data in their region or industry.

The primary purpose of this workshop was to present the initial case study findings, and sense check current thinking against the knowledge and understanding of experts in the field. Beyond this, the workshop aimed to explore what other regions are doing in relation to open data, and to consider the wider commercial opportunities presented by open data. The findings and ideas raised in the workshop have fed through into this summary report.

Seven breakout rooms were held over the course of the afternoon, focussing on the following key areas:

- **Data and data management**: what factors are most important when considering what datasets to publish?
- **Open data culture**: how can we develop a good open data culture?
- **The wider value of open data**: what does open data offer our cities and how can we measure its value?
- **Communication and awareness**: what makes a successful open data communications strategy?
- **Commercialisation:** which types of datasets would offer the greatest commercial potential?
- **Priority actions for Greater Manchester**: what does Greater Manchester need to do to create a better open data ecosystem?

- **Priority data for Greater Manchester**: what data should Greater Manchester make open?

Through these breakout rooms and discussions, the team was able to explore what other regions are doing in relation to open data and gain further insight into the commercial opportunities presented by open data.

The second part of the workshop built upon the first, and aimed to identify any early-stage findings about what Greater Manchester needs to do in order to open up more public sector data for re-use. Discussions surrounding the current and desired open data landscape in Greater Manchester have also fed into the overall findings in this report.

A summary of the workshop sessions will be released separately on the GMCA Local Data Review webpage.

What is this document?

This report concludes the analysis stage of the Local Data Review. It brings together all of the above strands of research, analysis and wider discussions into one summarised overview of our main findings. These findings are also complemented by discussion of the relevant national and local policy context.

The findings from this initial analysis draw out the core elements that an open data project **must** have, **should** have, or **could** have, in order to be successful, from a public sector perspective. This is relevant to Greater Manchester specifically, but many of the lessons learned are also valuable for other areas wishing to embark on a process of opening up more data for wider re-use.

Similarly, this review – and the wider Local Data Review – is concerned with the opening of data to support commercialisation and innovation, rather than from a democratic and transparency perspective. Nevertheless, the findings from the initial analysis can also be read with that approach in mind, and many of the findings are not specifically designed around how open data is used.

Section 1: Historical and strategic context

Historical overview (pre-2010)

Open data as a concept is not new. Open access to information in fields such as science and medicine has been a core concept for peer reviewing for many years. However, open data – in the sense of open access to government or public sector data – is a relatively modern conceptualisation, based on arguments of transparency and public funding. The opening of government – both national and local – in the UK gained initial popularity during the early- to mid-2000s, coinciding with the rising use of the internet, and movements to increase the free and open use of information, such as the creation of Creative Commons.

An important driver behind open government data in the UK at this time was the Guardian Newspaper's 2006 Free Our Data campaign. This called for the release of certain types of public data held by local authorities and government-funded agencies, like the Highways Agency and Ordnance Survey. The basic aim of this campaign was to persuade the government to abandon copyright on essential national data that had been collected using taxpayer's money, making it freely available to the public for re-use. Shortly after the launch of this campaign, Transport for London became one of the earliest public sector organisations to start providing some of its open data online in 2007.

Momentum for open data built over the subsequent few years. In 2009, the government released 'Putting the Frontline First: Smarter Government', which arguably made open data mainstream in government by setting out plans for the creation of a single data portal for public sector data, and the release of more open public data.

Between 2008 and the early 2010s, a number of open data initiatives and strategies were launched. Data.gov.uk, a single, easy-to-use, online access point for public data, was one of the most notable open data projects launched in this era. The website still continues to provide a space for government departments, councils, and other public bodies to share non-personal data openly. It was primarily established to build a greater level of trust among citizens, while also hoping to deliver cost-effective and better targeted services.

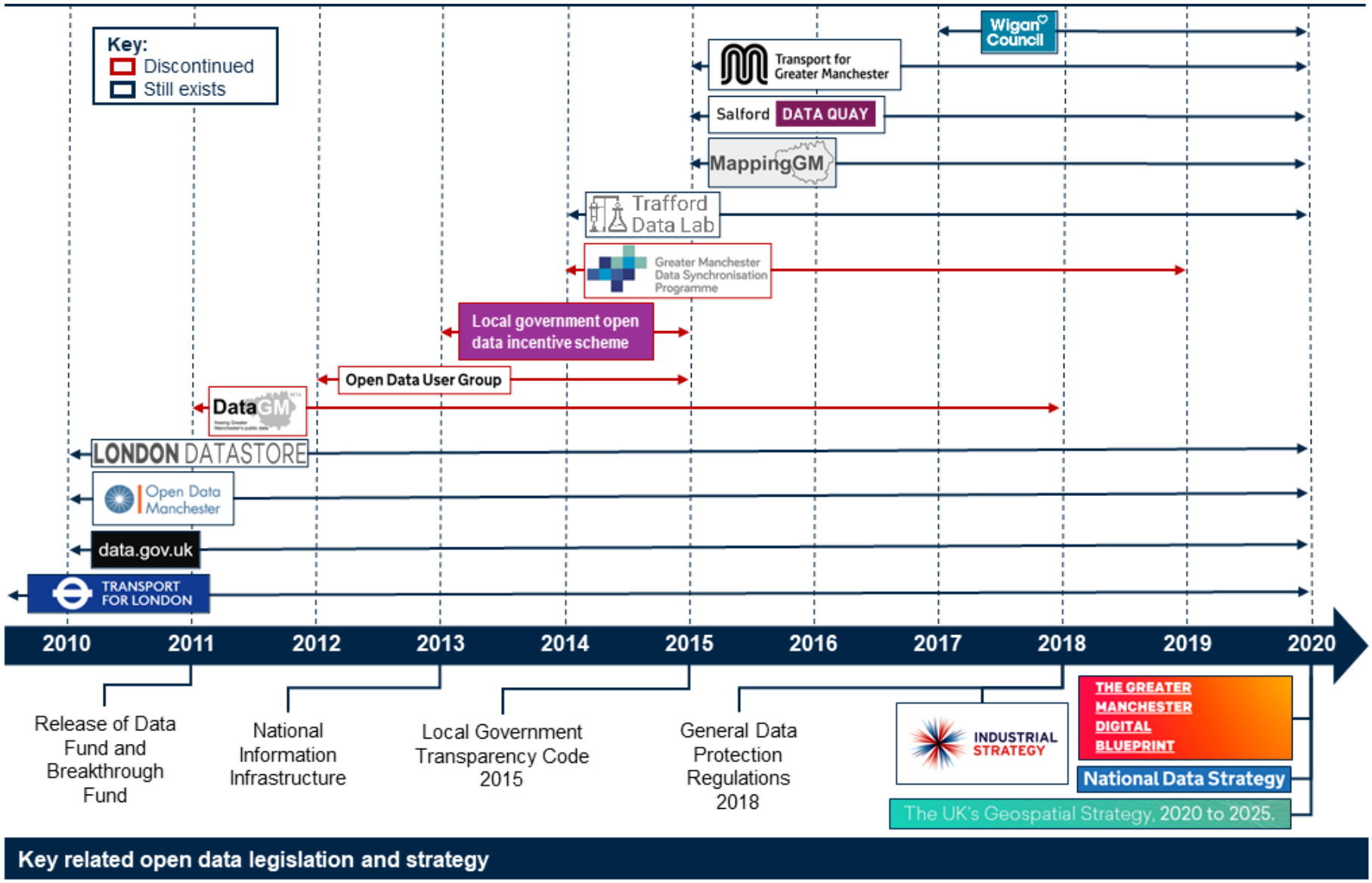
Alongside the creation of data.gov.uk the Open Government Licence was developed in 2010. The Open Government Licence was designed as a tool to enable public sector organisations to license the use and re-use of their data under a common, consistent, open licence.

At this time, the Cabinet Office set up the Public Sector Transparency Board (now called the Data Steering Group), to further drive forward the government's open data agenda. This board drew up a set of Public Data Principles, which pushed the open data agenda further, with a particular emphasis on government transparency. The Public Data Principles also outlined the need for public data to be released under the same open licence to enable free re-use, including commercial re-use, of the data.

Government transparency was the key driving force behind public sector open data throughout most of these early developments, with the development of better public services a secondary aim. Since 2010, however, open data in the public sector has expanded significantly in scope and objectives, both nationally and locally. Figure 1 below illustrates a selection of these expanding open data projects from 2010 to

2020, alongside some key open data legislation and strategy during this recent period. These programmes and strategies will be explored further in the subsequent section.

Figure 1: Timeline of selected local and national open data initiatives



Source: GMCA Research, 2020

National context - 2010 onwards

In the earlier half of this decade, the strategic push for open data still had a strong focus on government transparency and accountability. For example, the Local Government Transparency Code 2015 was developed largely to support the transparency agenda. This code sets out the minimum information that must be published by local government organisations, including details such as: items of expenditure over £500; procurement information; senior salaries; and annual audited financial statements. While the primary focus of these data are clearly around government transparency, the Transparency Code also recognised the wider potential of releasing these data. Specifically, it outlined how the availability of public data could also open new markets for local businesses, voluntary and community sectors, and social enterprises, which it expected could generate significant economic benefits.

These types of ideas around the broader value of open data have gained momentum and become more mainstream in the later half of the 2010s. While open data will always remain important from a transparency perspective, it is now generally understood to be an asset in itself. Utilising all government assets fully is ultimately key in terms of providing better public services, growing the local and national economy, supporting the creation of new jobs, and improving productivity. Open data is therefore now understood in terms of its potential to support these broader development objectives.

Recent government policy reflects this shift in strategic focus. For example, the recently published National Data Strategy (NDS) recognises that better data and better use of data can improve existing service delivery, drive scientific and technological innovation, open up new markets, and drive demand for a skilled workforce. In recognising the UK's role as a leading digital nation, the NDS explores how to leverage our existing strengths to further increase the use of data across businesses, government, civil society and individuals. The NDS builds upon existing initiatives like the UK Industrial Strategy, AI Review and Research and Development Roadmap. Together, these strategies and initiatives set out a framework for further investment in data to strengthen the national economy and create future opportunities.

The UK Industrial Strategy in particular is relevant for the background context to this report, given that our Local Industrial Strategy has emerged from this. The UK Industrial Strategy sets out a long-term plan to invest nationally in skills, industries and infrastructure, boosting productivity by creating good jobs and increasing the earning power of people across the country. As part of this broader objective, the strategy recognises that the UK holds a range of world-class data – such as the highest quality geospatial and climatic analysis, and high-quality, real-time transport information – much of which is held by public organisations. These data offer the potential for new products and services that could transform the UK economy and wider society, and so the strategy commits to making more of this data available to innovators and businesses across the country. As part of this commitment, it also highlights the importance of supporting businesses in accessing and using public data, given that data can only be seen as an asset if it is being used to add value. The measures set out in the Industrial Strategy aim to support the UK's tech industry to flourish, and reinforce the UK's position as one of the world's leading digital economies.

The broad aims in the NDS and the UK Industrial Strategy align with the aims of the Government's Geospatial Commission, which was set up to take advantage of the significant opportunities that geospatial data and technology offer to the UK. The Commission has recently negotiated a £1 billion investment in public sector geospatial data and has published an in-depth Geospatial Strategy, which aims to increase the amount of location data available to the public sector. In the longer term, this strategy hopes that improved location data will underpin the next generation of public service delivery, and support the growth and innovation of businesses across a range of sectors, from infrastructure to retail to the environment. In the shorter term, location data are an essential part of innovations and collaborations to manage the COVID-19 threat.

The high strategic priority assigned to open data on a national level throughout these different strategies and initiatives indicates the growing general awareness of its potential value to the UK economy. This national context provides a solid background on which to justify a strategic push for opening public sector data for re-use at a more local scale.

Local Context

The open data ecosystem in Greater Manchester can be best understood within the context of these broader national developments over the past decade. The national strategic framing of open data as an asset arguably supported the creation of DataGM and the Greater Manchester Data Synchronisation Programme, which were both attempts at creating a GM-wide datastore in the early 2010s. More recently, the emphasis on open data as an asset has led to a number of public sector bodies in the region – such as Trafford, Salford and Wigan councils – creating open data projects that release a range of different local data. These local data portals all go beyond simple transparency approaches to incorporate a wider range of data themes and analyses, ultimately adding value to local communities and businesses.

Alongside the development of these sophisticated open data programmes, open data is becoming embedded as a strategic priority within local policy, partly in response to the national strategic developments outlined above. In particular, our Local Industrial Strategy (LIS) echoes narratives from the national UK Industrial Strategy, but recognises that Greater Manchester in particular has a "fast-growth" opportunity in digital and data as a key enabling technology across all sectors. The LIS recognises that growth in this area should continually propel the city-region's economic performance by providing highly productive jobs and organisations, as well as driving productivity improvements and inward investment. This has clear parallels with the ideas and approach in the national Industrial Strategy.

In addition to the LIS, our Digital Blueprint aims to ensure that Greater Manchester becomes internationally recognised as a centre of digital innovation, research and practice, and has the digital infrastructure needed to become a world class digital city region. The Digital Blueprint recognises that opening up as much non-personal data as possible by default is key to meeting these broader objectives.

Through these local strategies and open data programmes, we are beginning to understand and articulate the true potential value of open data to the local economy.

Open Data as a 'Hype Cycle'

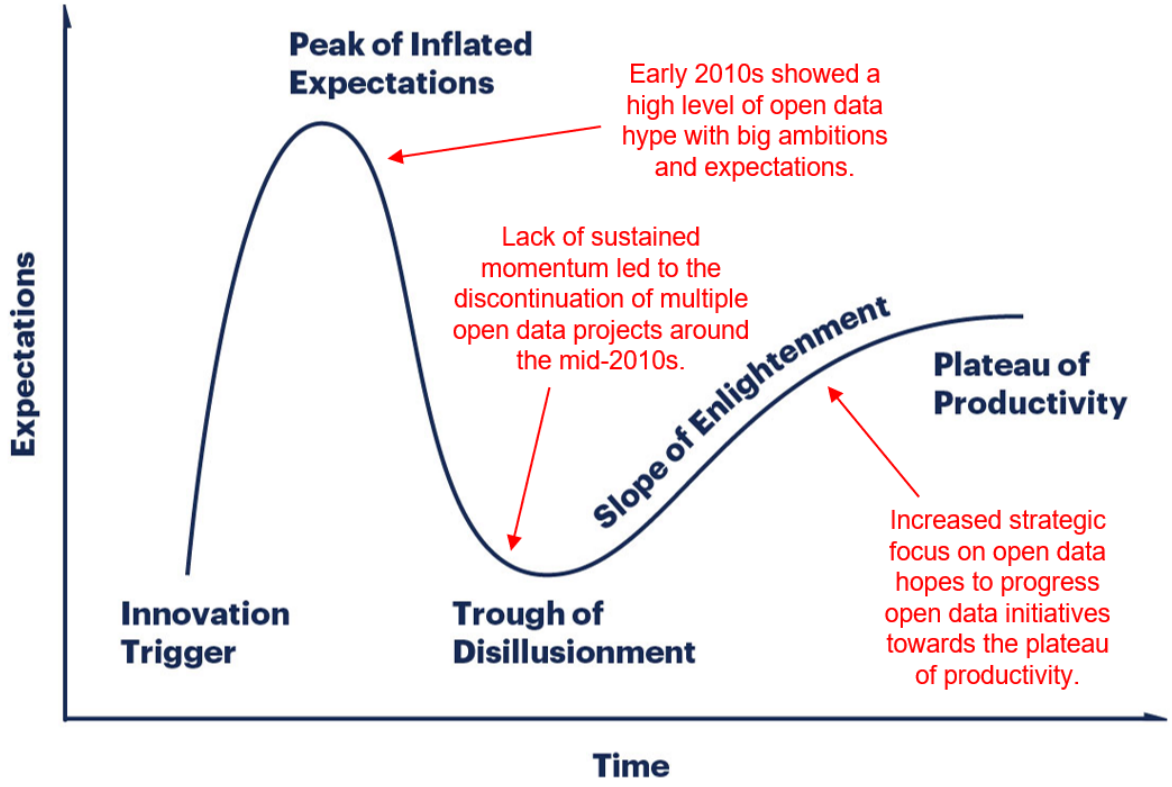
The timeline of open data programmes and developments shown in Figure 1 and discussed above broadly correlates to the various stages of the Gartner Hype Cycle shown in Figure 2, below. The Gartner Hype Cycle provides a graphic overview of common maturity and progression in the adoption of new technologies over time. The high enthusiasm for open data around 2010, as noted in the previous section, arguably lead to a 'peak of inflated expectations' at this time. Some of the national and local initiatives and organisations from this era, such as data.gov.uk and Open Data Manchester, have stood the test of time and are still around today, but others have proved less sustainable in the long term.

There were several open data initiatives during the mid-2010s that were ultimately discontinued. For example, the Open Data User Group – a national body that gathered the opinions of open data users to influence public sector data releases – operated between 2012 and 2015. This coincided with the Local Government Open Data Incentive Scheme, which aimed to publish local government data in a consistent open format to create a national view of local authority data. As with the Open Data User Group, this scheme only ran up to 2015.

While these projects had ambitious aims, the practical challenges of overcoming the fragmented and inconsistent approach to data across different local authorities and public bodies led to an arguable 'trough of disillusionment' in relation to open data in the early- to mid-2010s. This is partly borne out in the use of data.gov.uk, which did not meet initial expectations for user numbers or datasets. With a predominant focus of open data as part of the transparency agenda at this time, opening up and improving the quality of open data arguably became less of a strategic priority for many organisations. Anecdotal evidence from our case studies suggests these challenges may have been exacerbated by reduced funding availability in the early 2010s, leaving public sector agencies unable pursue open data initiatives.

More recently, however, there has been renewed interest in open data – as discussed above through the national strategic attention given to open data in the UK Industrial Strategy, the National Data Strategy and the Geospatial Strategy. Each of these strategies recognises the importance of open data as an asset for the UK economy, and have pushed the use and re-use of public sector data higher on the national agenda. With this renewed strategic focus on open data, and the commercial opportunities from data assets, the UK is arguably now entering the phase of the Gartner Hype Cycle focused on growing awareness and enhanced productivity from open data initiatives. Sustained strategic prioritisation of open data at the national and local level will be key to maintain this enthusiasm, which will ultimately allow the full potential of open government data to be realised.

Figure 2: Gartner Hype Cycle, with a focus on open data



Source: Gartner, with comments added by GMCA Research, 2020

These broad national trends in enthusiasm for open data across the 2010s were largely echoed on a local scale. The high enthusiasm in the early 2010s led to the creation of DataGM and the Greater Manchester Data Synchronisation Programme around the same time. These both had high expectations about what open data could deliver locally. Yet the practical realities and challenges faced by these projects did not live up to the heightened expectations of what they could deliver, and both projects were ultimately discontinued – correlating to the 'trough of disillusionment' experienced nationally at the same time.

There was a reduced availability of resources to pursue open data initiatives locally in the early-mid 2010s. At the same time, the rising importance of devolution over the same period created both a challenge and opportunity for open data – a challenge to stay relevant and impactful, and an opportunity to better align initiatives across the city reion.The history of co-operation amongst public sector organisations within Greater Manchester has laid a strong foundation on which to build and sustain regional collaboration in relation to future open data priorities.

Since the mid-2010s, the perceived dip in enthusiasm for open data on a local scale has arguably been surpassed by the rising awareness of open data's potential and opportunities. This can be seen throughout the rise in local open data initiatives and through the strategic prioritisation of open data within local strategies, as outlined above. This shift on a local scale is in line with, and likely a result of, the shift seen nationally – the local open data agenda is becoming less focused on transparency and more focused on supporting commercialisation and the use of data assets to support regional economic growth.

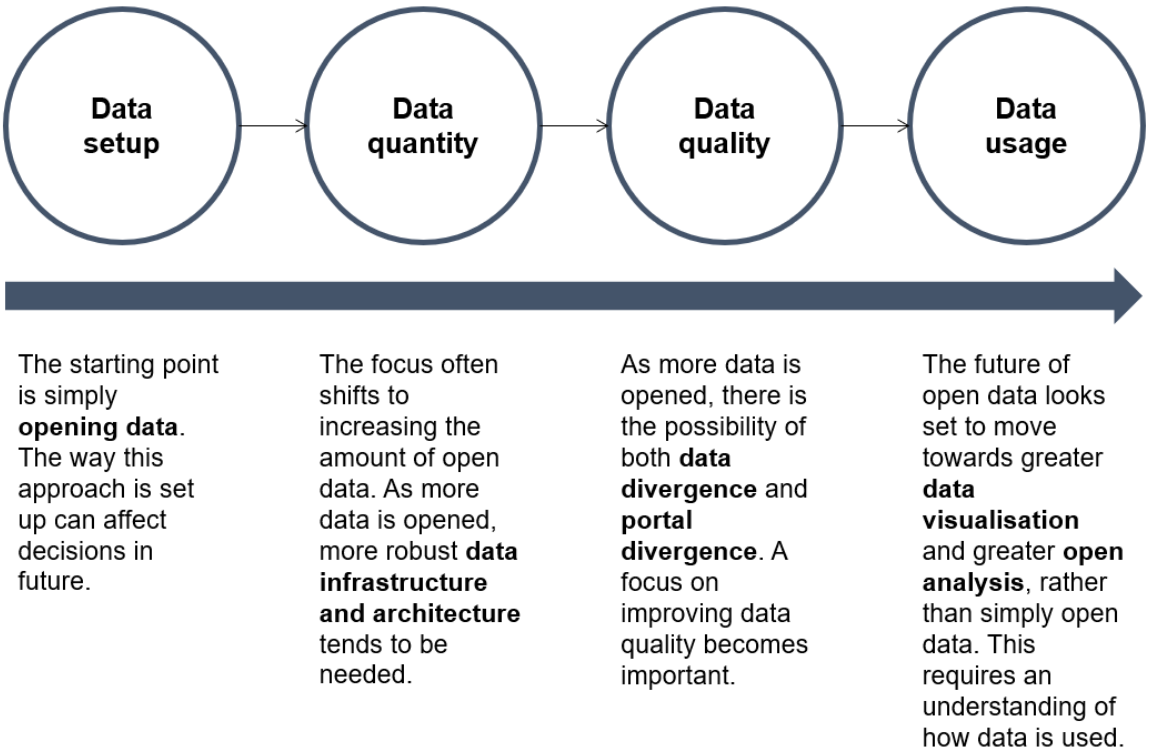
This report sits within the historical and strategic context outlined above and hopes to build further momentum for opening local public data – pushing the local open data ecosystem towards a ‘plateau of productivity’. Reaching this open data stage in the Gartner Hype Cycle ultimately has the potential to drive innovation and growth across the region.

Section 2: What have we learned so far?

Evolution of open data projects

Our analysis has identified a common four-stage evolution of open data projects and programmes. In short, once data has been opened, the focus generally moves from increasing the **quantity** of information, to increasing the **quality**, to improving **usage** of information. At each stage, a business case is usually made to progress the work forward. These broad stages and details of the key focus at each stage are summarised in Figure 3 below.

Figure 3: Open data project development timeline



Source: GMCA Research, 2020

These stages also equate to different periods of the Gartner Hype Cycle identified previously. That is, the move to increasing the *amount* of open data is similar to the increasing, and potentially falling, expectations that initially occur. In order to raise productivity and use of open data, the focus must instead switch to quality of open data, providing value to users. Finally, a focus on how users might use open data increases both the value and usability of information, allowing for an increase in productivity.

Key features of open data projects at each stage

With the observation of this general open data progression, it is possible to identify the important features at each stage. Not all projects and programmes move in the same way through this process. There may be jumps, or a more detailed focus on

one of these sections – for example, a focus on increasing data quantity. Nevertheless, to add value to the work, often projects will need to move on to the next stage of development. The main features at each of these stages are summarised below.

Stage 1: Data setup

The starting point is simply opening data. This requires basic key steps, such as: deciding what datasets to make open; having a place to put the data; creating a valid business case for opening data; and agreeing responsibility for uploading and maintaining datasets.

The data itself is the first and most important feature in any open data initiative. At this initial stage, setting out basic data standards and data principles are key to ensuring that any data opened is complete, accurate and usable. For example, one guiding approach to data principles was designed by the opengovdata.org project, which set out '8 Principles of Open Government Data'. These 8 principles outline that open data must be: complete; primary; timely; accessible; machine processable; non-discriminatory; non-proprietary; and license-free.

Once basic data standards and principles in place, alongside the key basic steps outlined above, the data setup stage is complete and basic open data is ready to be shared.

Stage 2: Data quantity

At this stage, most open data projects have progressed from sharing simple data files on a website, to holding a central repository for data. In Greater Manchester, some public sector organisations release data on different webpages, on a single webpage, or via a data portal. The data standards and data principles from the first stage remain key as the quantity of data released increases, and throughout the subsequent stages. These standards and principles should ensure that the expansion in data quantity does not come at the expense of data quality.

When focusing on increasing data quantity, open data programmes will need to think about what data assets they hold and which of these should be opened as a priority – often drawing up a pipeline of data releases. However, this is not always completed as a priority, and often the release of data is based on specific asks, or through what is currently available. This can create a challenge of ensuring that data can be released consistently, and that there is a set data owner responsible for the regular release of data.

As part of increasing data quantity, organisations may also seek to provide options to download data in multiple file formats. This makes data accessible to a wider audience and gives users a wider range of options to manipulate and interpret their data. The NOLAlytics data portal in New Orleans is a good example of this, allowing users to export data in CSV, RDF, RSS, TSV and XML formats.

As part of progressing to a specific website with an underlying datastore, thorough consideration of the required underlying digital infrastructure is also likely needed. Popular choices across the collection of case studies appear to be CKAN, Socrata, and DataPress (who use CKAN, but provide additional support) and AWS (usually with an open source solution), but other options are available too. The type of

websites or data portals created at this stage can range in size and scope but does not have to be expensive, as illustrated by the Salford Data Quay case study below.

Example 1: Salford Data Quay (SDQ)

After building up robust internal data architecture, Salford City Council launched the external open SDQ platform in 2015. The SDQ was set up as an open-source CKAN data portal with an initial grant of just £31,000 from central government. This portal has served the SDQ since then with minimum ongoing maintenance costs or adaptation of the basic design, illustrating that open data platforms need not be expensive.

Stage 3: Data quality

As data quantity increases, especially if these datasets are from multiple organisations, departments or teams, there is the possibility of both data divergence and portal divergence. The more organisations or datasets are involved, the more likely there will be non-standard datasets and variation in data quality. A decision needs to be made here between types of data portal – a centralised, single datastore with a greater focus on quality; or more of a federated data portal approach providing access to datasets held by different organisations, and with a focus on greater quantity, rather than quality. Each choice has pros and cons, as well as competing interests, as illustrated by the comparisons between the European Data Portal (EDP) and European Union Open Data Portal (EUODP) in Example 2 below.

Example 2: European Data Portal (EDP) and European Union Open Data Portal (EUODP)

The EDP and EUODP are both maintained by the Publications Office of the European Union. The EUODP is a centralised portal of around 15,000 datasets, exclusively provided by EU institutions and agencies. It is specifically created by and for the European Union. By contrast, the EDP is a federated data portal with over 1 million datasets from across the continent, including those from the EUODP but also from other EU member states and affiliated countries. This illustrates the first key difference of the two portal approaches: data quantity.

The second key (and related) difference is a matter of data quality. The more centralised approach of the EUODP leads to higher data quality and more consistent interoperability between datasets than the federated EDP portal. The large scale of the EDP creates a challenge in this regard, but has the advantage of allowing publishers from across the continent to more easily share datasets. Ultimately, portal design comes down to an arguable trade-off between quality and quantity, which must be balanced appropriately in relation to the primary objectives of the portal itself.

Whether a centralised or federated model is chosen, there are common features at this stage that help to maintain the focus on improving data quality. For example, having data co-ordinators and data leads embedded in senior leadership can help

with the oversight of data and ensure that data meets the expected standards and quality.

Building automation into the publication of new datasets is also common at this stage. For example, the Ordnance Survey has a rolling and continuous update process for its open data. This gives users assurance that they are always interacting with the most up to date and highest quality version of the data.

Stage 4: Data usage

At this more mature open data stage, the focus shifts to understanding and supporting the users of open data, and to driving up its usage. There are a wide range of approaches at this stage, with features like data visualisations, interactive maps, and dashboards and being commonplace across the collection of case studies. These features all enhance the accessibility and usability of datasets and encourage greater understanding and interpretation of data. The example of the London Datastore below illustrates just some of the many diverse potential options for data visualisations, and outlines how these features can support users and offer practical use.

Example 3: London Datastore

The London Datastore contains a wide range of interactive maps and infographics to visualise its data. For example:

- The London Infrastructure Map is an interactive tool to explore current and future development and infrastructure projects. It gives utilities, boroughs and developers a clear picture of what developments are taking place, so that they can plan better for new housing and other changes.
- London Area Profiles allow users to explore a dynamic view of different districts of the city with filters and layers that provide comprehensive insights into their socio-economic, demographic, health and education statistics.

These types of visualisations significantly enhance the usability of data, making data more accessible to both expert and non-expert users.

Some open data initiatives go one step beyond data visualisations, to actively encourage open analysis of their data. The example of Data Mill North below illustrates how these measures can encourage greater usage of open data in a way that adds practical value to the local area.

Example 4: Data Mill North

Data Mill North has creatively encouraged open analysis of its data through a range of techniques. For example, they hold open-ended Innovation Lab sessions to find data-led solutions to novel community challenges. These sessions ensure that existing datasets do not get overlooked and also identify new data uses.

The Leeds Bin Map is a useful tool that came from previous Innovation Labs. This app allows residents to register for automatic bin collection notifications for their

postcode, which has led to a reduction in missed bin collections by 50%. Furthermore, the information about who has accessed the app can theoretically be shared with the council to reduce the use of letters or mailouts to these properties.

Similarly, the Social Housing Picker was developed using the open data on Data Mill North. The tool used information about waiting times for social housing to allow individuals to understand the likely waiting times for social housing given their particular circumstances and location. Furthermore, it provided information as to areas nearby that might have shorter waiting times. The encouraged use and analysis of data released on the platform led to the creation of the tool, which aims to be updated based on new lettings category amendments from 2019.

While the move towards greater data visualisation, analysis and use is popular at this mature stage of open data, it is only one of many possible scenarios. While these ready-made analysis features tend to suit non-specialist users, different parts of the user base may have different needs and preferences. In particular, developers may prefer aggregated, raw data that can be easily accessed using application programming interfaces (APIs). An open data programme can opt to separate the front end portal into different sections to cater to multiple audiences, or can choose to focus more heavily on one part of the user base. For example, while the London Datastore has built data visualisation and analysis into its datastore, TfL and the Food Standards Agency have opted not to create public-facing apps or developments from their data. Instead, they focus on making their data open and high quality. The choices made here depend in part on the expected needs of the end users – there is value in both of these approaches.

Where is Greater Manchester on this journey?

The case studies, and feedback from our open data workshop, suggests that Greater Manchester is at varying stages in this progression, dependent on the organisation or area covered. Some organisations, including the GMCA, are at the first stage and largely open up the data that they are statutorily required to open. These include financial statements; expenditure above £500; and senior salaries.

Other organisations have moved onto stage 2: creating a central repository for data, and ensuring that all data is stored there. Some organisations have applied this only to certain types of information – for example, Wigan has created a data portal focused specifically on GIS information – while others have focused on a wide variety of information and sources.

Previously, Greater Manchester has attempted stage 3 and 4 projects through DataGM and the Greater Manchester Data Synchronisation Project (GMDSP). Neither of these projects ultimately proved sustainable – arguably due to many areas in the region not having levelled up to stage 2. However, they were very much ahead of their time – bringing together datasets locally for wider value and re-use, before this idea was mainstream on the local and national open data agenda.

In terms of basic approaches, it would appear that an overriding aim for Greater Manchester would be to develop a roadmap taking relevant organisations from stage 1 (opening data) through to stage 4 (improving data usage). In the first stage, this should include support to ensure that all relevant organisations are publishing

existing statutory datasets consistently. Building on from this, setting standards for data and metadata that are consistent across public sector organisations in Greater Manchester would lay the foundation for later attempts at stage 2 and 3 projects. These consistent standards would potentially also support the creation of a GM-wide data portal at some point down the line. Further along the roadmap, outlining strategies and approaches to build up an active user community, and setting out simple ways to create visualisations and open analysis, should help organisations achieve stage 4 objectives and improve data usage.

Perception and reality

Recurring challenges and concerns surrounding open data have been identified across our collection of case studies and voiced within our open data workshop. However, alongside identifying these challenges, we have also found evidence to suggest that few have created problems in practice. This highlights a key message that some barriers to open data are perceived rather than real. Some of the main commonly identified challenges are outlined below, along with some limited and anecdotal evidence to suggest that these may not have been as significant a challenge as first identified.

Misinterpretation and misuse of data

The potential misinterpretation and misuse of open data were identified as concerns in a number of case studies, as well as during the open data workshop. Some organisations voiced initial concerns surrounding the potential misinterpretation of their data by third parties and its potential to be used against the organisation. Despite this obvious concern about the way open data is used, none of our case studies highlighted any instances of data being used in a way that harmed the organisation. More importantly, this identifies a need to communicate early on that the organisation has no control over how that open data might be used once it is released.

Information governance

In a similar vein, some participants voiced concerns about information governance being a potential barrier to open data. For example, one workshop attendee identified a resistance from its information governance team to publish data due to privacy and legal concerns arising from recent GDPR regulations. Whilst this resistance was noted, it was not identified as having stopped the release of open data.

In practice, the only case study where GDPR or information governance had affected the release of open data was with data.gov.uk. In this instance, the Department for Education (DfE) had tightened their data release procedures, removing the assumption that any anonymised data should be published. This limited the potential number of datasets that could be released, but did not stop the opening of DfE's data outright.

DfE's experience appeared to be the exception rather than the norm, and none of our other case studies presented any examples of information governance acting as a significant barrier or challenge to opening data. Privacy concerns – especially with regards to non-personal data – are thus more likely to be perceived than real, indicating a need to provide practical reassurance on this front. Nevertheless, it is important that open data projects retain a focus on appropriate information governance, regardless of the type of data released.

Rather than seeing information governance as a challenge to releasing open data, Insight Cheshire East has used information governance requirements to drive up internal data standards and processes. Insight Cheshire East maintain and publish an Information Asset Register on their data portal, a complete list of 1,014 council information assets. Each asset is listed with an information rating, how the

information is stored, who is responsible for it, and a brief description. This register makes it easy for the Council to keep track of what personal information it holds to ensure GDPR compliance, and allows the Council to verify that it is meeting its obligations under the Re-use of Public Sector Information Regulations 2015. Not only does this register meet regulatory requirements, it also allows for quick identification and management of any risks to information held.

It is important to reiterate to both senior management and coleaues that GDPR and information governance is by no means incompatible with open data releases, as illustrated by Cheshire East Council's information governance approach. This message is important, as leadership support has been identified across a number of case studies as key to the success and long term sustainability of open data initiatives.

Lack of benefits

Discussions at our open data workshop highlighted that a big concern from the public sector perspective is how commercially successful open data is, and whether it will generate internal benefits or simply become "another tick-box exercise". This concern was also echoed throughout a number of case studies. However, it is important to recognise that open data can be hugely valuable without providing any direct financial gain to the host organisation. Open data, like transport infrastructure, is an economic enabler – it provides benefits downstream from those who actively share and interact with the data. A detailed report entitled Open data: Unlocking innovation and performance with liquid information by the McKinsey Global Institute (MGI) outlines this process more thoroughly. This report argues that capturing the market size of open data requires an understanding of the products and services that are improved or enabled through open data.

While a direct financial gain may or may not occur to the host organisations, open data is likely to provide a range of wider economic and strategic benefits, as outlined in the subsequent section on the value of open data. TfL's open data, for example, is not primarily aimed at having a direct impact on the organisations – or its transparency. Its open data is aimed at developers who specifically aim to use the data in their products and services. These economic and strategic benefits, as well as the wider types of impacts outlined by the MGI report, can be outlined in the business case for open data projects, as part of the wider economic value of open data projects.

Trust and privacy

Some participants in the open data workshop raised trust and privacy as a key concern with certain types of open public data, particularly in relation to sensitive personal or human movement data. There was a broad feeling among some members of the group that the general public does not trust the parts of the public sector with their data, and many are sceptical of personal data being used for commercial benefit.

These sentiments were supported by a 2018 Eventure Research survey into public attitudes to data and information sharing for public benefit, for the GMCA. This survey found that respondents generally reported lower levels of trust in local

authorities using data appropriately than other parts of the public sector, such as the NHS, Police, and Fire and Rescue services. In particular, there was scepticism around personal data being shared from the public sector to the private sector, with most focus group participants holding negative perceptions of the private sector in relation to its use of personal information.

Despite these challenges in public perception, almost three quarters of participants (74%) agreed that they would be more confident about public sector bodies sharing personal data and information if they could see who their information is shared with. Participants could also see potential benefits to sharing data in terms of efficiency savings for taxpayers and in terms of individuals receiving joined-up services. This implies that negative public perceptions in relation to trust and privacy are not fixed, and could be improved with better transparency and communication of benefits. Further work to understand whether trust issues extend to non-personal information would be useful.

One participant in our workshop noted that Wi-Fi tracking on the London Underground in response to COVID-19 has actually been well received on the whole by the public, despite the potential for privacy concerns in relation to this type of data. TfL had already been tracking WiFi connection data to understand patterns of movement across the London Underground network over the past year. Their approach does not track or identify any specific individuals but provides useful aggregated insight into where the network gets crowded, at what times, and how this changes in response to events and network alterations. This data is shared in near real time on the TfL website, allowing customers to better plan their journeys and avoid congestion or disruption. During the COVID-19 pandemic, this technology is helping to monitor social distancing and identify high risk areas, and is allowing customers to make informed decisions about the safety of their travel.

The broad public acceptance of WiFi tracking on the London Underground illustrates that opening this type of data can be highly successful – in this case, offering value from both a customer perspective and a broader public health perspective. The key is to clearly communicate the benefits of sharing and using this type of data, and to communicate how opening certain data is primarily in the public interest, rather than primarily for commercial gain by the council. It is also worth noting that most open data relates to non-sensitive and non-personal information, where there are likely to be fewer concerns from the public around opening and sharing this information. Therefore, while fears around trust and privacy are certainly important considerations for any open data initiative, this need not necessarily transpire into a barrier for opening public data.

The value of open data

It is fundamentally important to recognise that the value of data is in its use, not in its storage. This means that open data is not valuable in and of itself – it must be used to have value. The case studies and workshop identified that not all open data has been used; in several cases, there were examples of datasets opened for a considerable time that had never been used. With the value of open data intrinsically linked to its usage, aims to increase productivity and economic growth through opening data must also look to increase usage, as outlined through some of the suggested open data approaches in Section 4 of this report.

However, where open data is being widely used, and used effectively, it can offer a wide range of different social, economic, environmental and other benefits. This section focuses specifically on: the economic value of open data, given the broader economic objectives of our Local Industrial Strategy; and the strategic value of open data, given the strategic context of our Greater Manchester Digital Blueprint.

Economic value

There have been numerous attempts to quantify the economic value of open data by several different organisations at several different geographical scales. On an international scale, the European Data Portal has been particularly active in this regard, and regularly undertakes assessments on the value of open data in the European Union. The European Data Portal's most recent estimate placed the market size of open data across the European Union at €184.45 billion in 2019, supporting 1.09 million open data employees and creating a whole host of wider reaching benefits. Some of these open data economic benefits are directly observable in terms of cost savings, increased revenues and Gross Value Added. For example, there may be cost savings for organisations in simply opening their own data, cost savings by external organisations in terms of acquiring open data for free or at minimal cost, or cost savings through efficiencies enabled by the re-use of open data.

Figure 4: Estimates of the value of open data

€184.45 billion

Estimated size of the open data market in the European Union in 2019

Source: The Economic Impact of Open Data, European Data Portal, 2020

27 million

Etimated number of hours saved on public transport by users across the European Union as a result of open data in 2019

Source: The Economic Impact of Open Data, European Data Portal, 2020

£130 million

Estimated annual economic benefits generated by TfL's open transport data

Source: Assessing the value of TfL's open data and digital partnerships, Deloitte, 2017

As well as noting these directly observable economic benefits, the same European Data Portal report was also keen to emphasise the wider secondary and indirect benefits of open data. For example, the report noted that open data can: generate time savings for the users of data; support knowledge economy growth; support potential new goods and services; increase the efficiency of public services; or encourage the growth of related markets. With such a broad range of potential secondary benefits, it can be hard to directly quantify the overall economic value of open data. This highlights how, similar to infrastructure, open data can be seen an economic enabler – it makes vital contributions to enable a "growing, innovative, ethical" economy. It is important to understand these broader indirect benefits as well as the more obvious direct benefits of open data – doing so provides momentum for the open data agenda and an evidence base to support a valid open data business case.

Studies on a more local scale have further supported the case for open data as an asset and an economic enabler. For example, our Transport for London (TfL) case study identified a range of internal and external economic benefits generated through TfL's open data. From an internal perspective, the tools produced by external developers have reduced TfL's internal costs by avoiding the need for internal application development. This has reduced the cost of SMS transport tracking services and reduced contact centre call volumes. Yet from an external – and arguably more significant – perspective, Deloitte recently estimated that TfL's open transport data generates £130 million annual economic benefits for travellers, London and TfL itself, supporting London's broader economic agenda.

These studies by Deloitte and the European Data Portal provide further support for the open data ambitions within the Greater Manchester Local Industrial Strategy, which recognises that opening public sector data for re-use has the potential to drive innovation and economic growth across the Greater Manchester region.

Strategic value

Beyond the economic and commercial value outlined above, open data can hold huge strategic value to host organisations and regions. Successfully opening up more data and enhancing open data analytics sends a wider strategic message about being an open and innovative place. This message can ultimately drive inward investment and attract significant attention from external regions and investors, whilst also supporting and retaining existing investors and businesses. The Australia and New Zealand Infrastructure Pipeline (ANZIP) is one example of where open data has been used to send a wider strategic message to drive inwards investment.

Example 5: Australia and New Zealand Infrastructure Pipeline
The ANZIP portal provides a forward view of major public infrastructure activity across Australia and New Zealand. Its creation signals a joint commitment to Trans-Tasman collaboration, open markets, innovation and investment by building a more integrated infrastructure market between the two countries. The portal provides details on a range of infrastructure activities, including the project status (from 'Prospective pipeline' to 'Recently closed), the value of the project, location of the project and wider contextual information.

By providing this information, this initiative provides opportunities for both investors and potential work bidders, but ultimately sends a high-level strategic message to private investors that Australia and New Zealand is a region worth investing in, and has attracted further investment in the region.

In the case of Greater Manchester, the potential strategic gains from opening more of our data clearly support the aims of the Digital Blueprint. Greater Manchester wants to be an internationally recognised centre of digital innovation, research and practice. Opening up as much non-personal data as possible by default is ultimately key to supporting this aim, and would drive the evolution and enhancement of the region's digital ecosystem. This could send a wider message that Greater Manchester is an attractive place to develop and grow – contributing to digital innovation, and boosting the region's external digital reputation.

Valuable data assets

The boxes below summarise some key data assets that were recognised as being valuable or widely used across the collection of case studies and at the workshop. Value in this sense includes valuable both from a commercial perspective, in that the data either drives decisions, supports innovation, or forms part of a product or service; and from a public service perspective, supporting the provision of better, and better targeted, public services.

It is important to recognise that being widely used does not necessarily equate to being valuable. Open data arguably becomes valuable when it is used for specific purposes and objectives – for example, to support decision-making, service delivery, risk assessment, or innovation. Each of the data assets identified below have an explanation of how they potentially provide wider value to the community of data users.

The datasets are not presented in priority order. They represent specific datasets or general data themes that would appear to have the highest demand for re-use from our case studies and workshops. The next stage of our analysis – working with businesses, academia and the third sector to understand their data needs – will seek to identify the detail of these datasets, and explore whether there are further datasets that we are not aware of that we should prioritise.

Priority Data: Transport

Transport data, from timetables and bus stops to real time information, were frequently cited as highly valuable within both the case studies and workshop. For example, TfL openly shares hundreds of datasets about public transport, roads, cycling and walking. TfL's open data powers over 600 travel apps in the UK and facilitates the development of technology enterprises. Up to 42% of Londoners are thought to use TfL's open data to make better and more informed transport decisions.

Open transport data therefore provide value to businesses and developers who can build apps from the data, but also provide value to anyone who uses these data to inform their transport decisions. The high popularity of these data suggests that datasets with a practical usage may have the greatest value to end users.

In Greater Manchester, TfGM is the public body responsible for co-ordinating the transport network and services. TfGM is committed to travel information open data and currently has an open portal that shares a range of datasets. These include Metrolink real time data, incidents and accidents data, car park locations and capacity information, and traffic signal locations and flow information. As well as this, TfGM is currently creating more open data feeds that are first tested and used internally, such as real-time bus and Metrolink information and service timetables.

Priority Data: Business rates

Business rates data was popular across a number of case studies. Business rates releases have no obvious direct internal benefit beyond increasing transparency, but provide huge value for the business community.

These data support market research activity and allow easier identification of potential local business opportunities, as well as providing basic contact information for existing businesses.

Much of Greater Manchester's business rates data is already released, either directly on local authorities' own websites, or through external datastores like data.gov.uk. However, not all of our business rates data is currently released and there would be value in fully opening these data across the full region.

Furthermore, while Data Mill North has attempted to standardise business rates information using a schema, the same approach has not been applied across Greater Manchester. Improving the standardisation of valuable datasets across the city region – especially when they are not produced by one organisation – will provide greater benefit to the organisations using that data.

Priority Data: Socio-economics and demographics

Some case studies, such as the London Datastore, found that local area profiles with an overview of health, education, demographics, and population projection data were very popular. This type of data, and the population projection data in particular, has a practical usage, as it is often relied on for forward planning and investment decisions by public, voluntary and private organisations.

Many of these socio-economic and demographic datasets are already released by the Office for National Statistics. However, there is some potential value in cutting and repackaging this data to a local level, for local users, in a format that is easier for the end user to consume. Wigan Council Open Data is a good example of where this has been done on a local scale in Greater Manchester – their Community Health Story Map combines several datasets together to give an overview of health across the district with a narrative explanation of what the health measures mean. This could have a wide range of potential uses. For example, the 'Loneliness Prevalence in over 65s' tab of this story map could help direct support services from voluntary organisations to residents in these areas.

Collaborating amongst local public sector organisations to produce these types of data consistently, and once for all of Greater Manchester, can both aid productivity within organisations and allow for greater comparability of information across the city region.

Priority Data: Planning

Planning data, such as the Brownfield Land Reister, planning applications, and Strategic Housing Land Availability Assessment (SHLAA) data, featured in many case studies. For example, the SHLAA map, showing where future housing development is likely to occur, was noted as the most popular dataset by Wigan Council Open Data.

These types of planning data have huge value to the private sector by identifying potential sites for development, or redevelopment, which can drive investment and growth in both the housing and commercial sectors. It can also aid investment decisions by utilities and other providers, allowing forward planning of activities.

Efforts have been made to standardise and open certain planning data across Greater Manchester through the Greater Manchester Spatial Framework. As part of this framework, the 10 Greater Manchester authorities have worked together to identify how they will collectively use land for homes, employment and infrastructure, with the aim of spreading housing and employment opportunities more evenly across the region. Opening this planning data in a consistent format across the ten local authorities aims to boost the prosperity of Greater Manchester as a whole.

However, while planning application data is released in Greater Manchester, not all local authorities provide geolocated information openly, or a single download of planning application data. Furthermore, a standardised schema – taking account of the variant planning policies between districts – could potentially be of great benefit, and allow planning application data to be more readily consumed. The Greater London Authority has taken steps to create a standardised schema, and a similar approach in Greater Manchester could provide a valuable, data-rich resource.

Priority Data: Movement and activity

In comparison to the transport data mentioned above,which is about the provision of transport services and accessibility, movement data specifically focuses on human usage of these systems. This can include information such as tram, train or bus usage; cycle path and road usage; traffic information, including density; and footfall data within towns and visitor attractions.

Human movement data can be collected via a range of methods, such as mobile phone tracking, WiFi counting, camera image counting, ticket purchases, and traffic sensors. These data can then be aggregated to build a more complete picture of travel and movement across an area. Greater Manchester benefits from having a single transport function to aid this collection and release of data.

These data are highly valuable to third parties, including sectors such as retail, transport, and healthcare. For example, retailers can use these data to better understand where to locate stores or run marketing campaigns; transport providers can use these data to increase the efficiency of their services and better manage demand; and healthcare services can use the information to direct patients with lung conditions away from more congested or traffic-prone areas.

Currently, not much human movement data is openly released in Greater Manchester, though it is actively collected – as mentioned above, through mobile phone signals, footfall counters or ticket sales. In several cases the information may be held by non-public sector organisations. However, there may be significant opportunities for the collection and opening of human movement data in the city region in the future, such as Metrolink ticket sales and footfall data.

Priority Data: COVID-19

Since the COVID-19 outbreak, many organisations have published COVID-19 data and have found this to be vital to both responding to the outbreak, and responding to the related economic impact. Examples include standard data released by central government on cases, hospital usage, testing and deaths; and more detailed pavement width information, providing detail on where social distancing might be more difficult. This data encourages informed decision making by both businesses and individuals, by building an understanding of the situation in an area and of any enhanced local risks. It is important to recognise that this type of data is only useful when it is up to date, given that the COVID-19 landscape can change fairly quickly.

COVID-19 data is not limited to information about the prevalence of the disease or the healthcare response. Many organisations from across the public, private, academic and third sectors are also interested in a variety of societal and economic indicators. This may include more detailed and up-to-date business performance information; licensing data; and non-domestic rates information. It may also include more detailed open data around services affected, such as bin collection data. An example of this information, related to the movement data above, is Leeds' use of Smart Bins, which monitor usage. Exploring the usage of Smart Bins over time indicates how COVID-19 has impacted town and city centre usage, and also shows when and where places are being used.

Priority Data: Asset location

Discussions at the workshop identified some potential benefits of opening basic asset location data held by public sector organisations. The opportunities and benefits here are broad. For instance:

- Sharing the location of electric vehicle charging points could help support the green agenda and encourage wider use of electric cars over polluting cars. While much of this data is crowdsourced and shared via applications like Zap-Map, there is the option of identifying potential new sites for charing points.
- Sharing the location of street furniture and benches could encourage citizens to spend more time outdoors, supporting mental health and social wellbeing.
- Street furniture such as lamp posts could also support the rollout of 5G, tying into the Digital Infrastructure priorities in our Greater Manchester Digital Blueprint;

- Detailed geospatial information about land and property ownership could help identify sites for infrastructure or property development, and where public sector sites could either support environmental or development initiatives.

Many of the standard street furniture assets, pathways, and roads are within a local authority's ownership. Collating single datasets of these in a consistent data schema can help to support the wider use of these datasets to support economic, health and social wellbeing.

Priority Data: Infrastructure and risk management

Several case studies had a specific focus on infrastructure. For example, ANZIP maps major infrastructure activity, such as construction projects over AUD$300m and investable greenfield and brownfield sites over AUD$100m, from planning through to completion. This type of data has practical usage and value for infrastructure investors and contractors and allows the market to prepare the financial and human resources needed.

Similarly, the Geospatial Commission and Cabinet Office are currently running a project on underground assets and reviewing strike avoidances, whereby one utility company encounters another utility company's assets by chance. Collating and opening data on underground infrastructure, such as electricity or digital networks, could more effectively support the co-ordination of future improvements and reinforcements to these networks.

Beyond this, opening infrastructure data in Greater Manchester could support the delivery of the upcoming Greater Manchester Infrastructure Strategy, which is underpinned by the existing Greater Manchester Infrastructure Framework 2040. This framework aims to develop and maintain a holistic infrastructure system that is robust and accommodates sustainable growth, which requires investment across all sectors from energy to transport to green and blue infrastructure. Sharing data on these infrastructure ambitions and their progression could simultaneously support effective and joined-up thinking, and drive inwards investment into the region to support the wider strategic objectives around infrastructure.

The potential use of these data suggests other datasets that enable risk management, for instance around flood risk indicators, may also be highly valuable to organisations in Greater Manchester. Flood risk data is already open, but indicators such as incidence reporting may be of benefit in terms of assessing broader flood risk and using this to inform local and commercial decision making.

There are other types of public sector data that have commercial potential but have not been listed above, as they are generally less relevant in the context of the Local Data Review. Health data in particular has significant commercial potential, but most of this is personal information that would be difficult to bring forward as open data. However, it is worth noting that efforts are already being made to more effectively link up health data with local authority care data in Greater Manchester through the Greater Manchester Digital Platform. This platform brings together data from multiple

sources and providers to provide a better and more joined-up service to those receiving health and care services. There is clearly huge economic and social value in this approach, but this involves better sharing of data within the public sector itself, whereas this report is primarily concerned with better sharing of data from the public sector to the private sector.

Additionally, certain regulatory data can have notable commercial potential – open regulatory data can help businesses to plan more effectively and minimise possible burdens, as illustrated in the Food Standards Agency case study. Despite this commercial potential, regulatory data has not been highlighted as priority data here. This is primarily because much of this data is controlled by national agencies, such as the Care Quality Commission or Ofsted, who often have the responsibility to collect their own data. Furthemore, when regulatory data is collected locally, or collected by local organisations – such as around food standards – it often remains the responsibility of the relevant national body to open the data.

Section 3: Challenges and barriers

Our research suggests there are a number of recurring challenges that are routinely faced by organisations when trying to open data. This section is by no means an exhaustive list of all of these challenges, but outlines the main recurring issues identified through the case studies, workshop and wider research to date. Mitigations to some of these key challenges are then discussed in the subsequent section.

Broadly speaking, challenges and barriers to open data can fall into two main categories: those that limit or prevent the initial opening of new datasets, and those that limit the long-term sustainability of projects once created. Most of the perceived barriers in the previous section arguably relate to fears surrounding the initial opening of data itself. However, this section covers a wider range of practical barriers that both prevent data being opened in the first place, and prevent its long term sustainability once opened.

Summarised findings and key points are presented in the 'Research notes' boxes for each of these identified challenge areas. Further detail is then presented in the subsequent text below.

Long-term sustainability

Research notes
A key factor in the long-term unsustainability of **DataGM** was its reliance on **short-term grants** as its main funding source, which created a sense of **temporality** amongst both developers and users of the site. **MyStockport** required large manual input to maintain and update the site, yet had relatively low usage. This **lack of value for money** led to the programme becoming unsustainable in the long-term. Open data projects need to **provide value for money – low cost or low maintenance over time, whilst having a direct and obvious benefit.** This can have a positive cyclical effect, creating a sense of permanence.

Sustainability is often the main challenge for an open data initiative, and programmes can ultimately become unsustainable as a result of a wide range of different challenges. Finances can be a key issue for open data sustainability, with some open data projects operating on the basis of short-term grant funding rather than through sustained internal funding, which can prove unsustainable in the long-term. This was the case for DataGM, which was not able to obtain a sustained and long-term funding agreement when its initial grant funding expired. This led to its eventually winding down in 2018. By operating on this short-term funding model, both the developers and users of DataGM did not feel able to rely on it becoming a permanent feature of the local data landscape. This sense of impermanence arguably limited the datastore's usage, given that many users would not want to become reliant on data from a source that they perceived as temporary.

The problems of short-term funding are therefore twofold: firstly, it can result in the discontinuation of an open data project when initial grant funding runs out; and secondly, this lack of long-term funding can create a sense of temporality that limits

user engagement with open data, reducing its wider value. A lack of long-term funding can therefore pose a huge barrier to the long-term sustainability and success of open data programmes.

Beyond funding, building and maintaining sufficient internal capacity and resources for running open data programmes is another key sustainability challenge. For example, in the early 2010s Stockport Council created the MyStockport platform – a simple datastore providing a range of key demographic metrics and service details – with the desire to help residents navigate council services and other aspects of their local area. The platform was set up when significant capacity and funding was available to dedicate to the project. Over time, however, the datastore became too complex and difficult to manage, requiring significant internal time and resources to maintain and update. Alongside the high financial cost and human resources needed to continue the project, usage was thought to be relatively low – ultimately, this was not an efficient use of public resources and the datastore did not offer value for money. Where the owners and users of open data do not feel that an open data project provides value for money, it is unlikely to be sustainable in the long-term. As such, the sustainability of a programme is dependent on its funding (with an aim of costs being low), capacity (again, with an aim of capacity usage being low), and value for money.

Leadership and culture

Research notes

The **Greater Manchester Data Synchronisation Programme** arguably **did not gain long-term leadership support**, which meant open data could not become fully embedded within organisational culture. This was a key factor in the eventual discontinuation of the programme.

Framing ‘open data culture’ around **transparency** misses the wider economic value of open data and **limits the opportunities for a valid business case**.

Open data projects need strong – and continuing – senior leadership, supported by an open culture. These can help to capture the wider, indirect benefits of opening data.

It can be challenging to embed a culture of open data within an organisation and gain sufficient buy-in from leaders to drive the open agenda forwards. The previous section of this report outlined a number of common perceptions surrounding open data, which have rarely turned into practical issues in reality. Yet our case studies sugest these perceptions often persist as common fears amongst strategic leaders and data managers, and can remain a barrier to opening data.

Even when initial fears and misperceptions are overcome, it can still be challenging to fully embed an open data culture within an organisation. ‘Open data culture’ is often framed in relation to the transparency agenda or in terms of internal benefits, such as reducing the number of Freedom of Information requests. This is a narrow vision that misses the broader purpose and value of open data. With such a narrow framing of open data, drawing up a valid business case can prove difficult as it can

be hard to articulate the full benefits of opening data, especially as they will not always directly accrue to the organisation opening the data.

Strategic buy-in from senior leadership is essential to move the open data agenda beyond this narrow focus on transparency and embed a broader understanding of the value and purpose of open data within organisational culture. This necessary buy-in can sometimes be limited. Where this is the case, data managers can find it hard to build momentum for open data projects – as was the case for the Greater Manchester Data Synchronisation Portal. While several local public sector employees were enthusiastic about getting involved in the project, a lack of long-term leadership support failed to embed the programme into organisational culture within the organisations. This meant that engagement with the programme usually fell outside of these individuals' work remits, so individuals had to find time and capacity to push the programme beyond their formal remits. This lack of long-term leadership support was a barrier that made the project as a whole unsustainable in the long term.

Even where leadership is fully committed to the open data agenda, a top-down approach to open data is unlikely to be sufficient on its own for encouraging wider cultural change. Some participants at the workshop noted that opening up data and embedding this within organisational culture is a complex process that requires flexible and creative thinking across all levels of an organisation, with buy-in from senior management to junior staff. Maintaining this buy-in and enthusiasm in the long term can be a great challenge (see the Gartner Hype Cycle in Section 1 for further information on enthusiasm and momentum in open data programmes). If efforts are not made to continually manage expectations and fully embed open data within the culture of an organisation, momentum can dwindle and projects can ultimately be discontinued.

Data and data management

Research notes

DataGM's data **overlapped** with the more widely used data.gov.uk. This **duplicated effort,** thereby **reducing engagement** with the datastore by some local organisations.

The European Data Portal attempted to instigate a **Quality Assurance Framework for metadata**, as an attempt to **improve standards and comparability of data**. Despite this, few datasets have achieved a rating of 'good'.

Simply having a **standards or metadata standards framework is not enough** in itself to guarantee that the framework is applied in practice, or to guarantee adequate data quality. There is a **tradeoff between data provider engagement** (which is also influenced by other factors) **and the standards applied to open data**.

Challenges around data selection, data quality and data upkeep can all become barriers to successfully opening data. Case studies like DataGM reveal the dangers of not addressing data selection challenges fully. In this case, too little consideration

was given to the gaps in existing provision of open data and to what data needed to be published. This approach – referred to as a “lob it over the fence” mentality in our workshop – meant that much of the data on DataGM overlapped with that released on data.gov.uk, a much more coherent, advanced and nationally reconised open data project. Many local authorities were already publishing data on data.gov.uk, and were therefore reluctant to duplicate effort and publish data in both places. This limited the datastore’s potential as a GM-wide data initiative, as not all of the public sector bodies in the region were on board with the project. DataGM had essentially launched a new open data programme without the necessary internal groundwork in place to scope out what data ought to be published. This proved ineffective and unsustainable in the long term, alongside the financial limitations noted previously.

As well as data selection challenges, achieving high enough standards of data quality can be difficult, particularly where data is shared from multiple sources and organisations. Arguably, the larger the number of organisations involved, the harder it can be to maintain standards. The stricter the imposition of standards, the lower the potential level of engagement; and the lower the standards, the higher the level of engagement, with a tradeoff of lower quality data or lower comparability.

This challenge was exemplified through the European Data Portal (EDP) case study. The EDP shares datasets from a vast array of organisations across the continent and has noted huge variations in data standards on its portal. These variations in standards can limit interoperability between datasets and impact data usability. Given that open data is only valuable when it is being used, data quality issues that impact data usability therefore limit the overall success and value of an open data initiative.

There are ways to mitigate this risk, such as creating a standards framework or metadata standards framework for publishing organisations. However, these mitigations also come with their own set of challenges and limitations. For example, the EDP created a Metadata Quality Assurance Framework, yet very few datasets on the EDP portal have achieved a score of ‘Good’ in this framework. This shows that simply having a framework in place is not enough to guarantee high standards and adequate data quality.

Beyond data selection and data quality, data upkeep can be a huge challenge to the long-term sustainability of an open data programme. Deciding what data to publish is the first and most obvious challenge in relation to open data, but decisions around what data to take down and stop publishing are also vital – and are often forgotten. If data publication is seen as the end goal and consideration is not given to the subsequent use of published data, time can be wasted continually updating and publishing datasets that have no wider value or use. This challenge means that data management can become timely and expensive when not done well, particularly if automation is not built into the process, and large manual input is required to update and maintain data. As noted previously, the resource required to maintain MyStockport, alongside its relatively limited usage, became both a financial and administrative burden over time that led to the eventual termination of the product.

Skills

Research notes
Many **SMEs did not have the necessary skills** to engage with the complex datasets offered by the **Copenhagen City Data Exchange**. This limited the site's usage and so the site failed to reach the critical mass of datasets and data purchasers needed to become sustainable. The GM Data Synchronisation Programme provided a **linked data platform that was ahead of its time**. As a result, the skills needed to engage with the data – both from a provider and user perspective – were not widely available. **Data must be provided to the skill level of the user.** It may therefore be relevant to **support the wider development of data skills**, internally and externally.

Skills can be a barrier to opening data, both internally and externally. From an internal perspective, some public sector organisations may struggle to build and maintain the skills to ensure that data is accurate, complete and up to date; or to develop and maintain the digital skills needed to build and manage a data portal. Some discussions in our wider research noted that anecdotally high rates of staff turnover within the public sector can sometimes prove a challenge in this regard.

From an external perspective, the case studies identified that there can be a lack of technical skills in the user base to fully engage with open data, particularly among SMEs. This lack of skills can make it hard for organisations to transform raw data into useful information, which limits the wider use of open data and therefore limits its overall value. This was the case for the Copenhagen City Data Exchange, which found that many SMEs did not have the capacity, budgets or skills to make use of the complex datasets offered by the portal. This limited the wider engagement with the site and arguably contributed to its eventual discontinuation.

Participants at the open data workshop built on this idea and noted that APIs can intimidate smaller organisations and less technical users, who may have fewer data skills and therefore prefer simplified data formats like a standard Excel spreadsheet. Open data initiatives that include too many complex datasets, or rely primarily on APIs, can therefore digitally exclude potential users who are not comfortable using this technology. The problems of this are twofold: firstly, this has a distributional impact and directs the benefits of open data towards organisations that have invested in these skills; and secondly, this model can prove unsustainable in the long term. Demand for open data is likely to be lower than expected if the potential user base does not have the adequate skills to access and engage with the data. Where demand is too low, open data programmes will not offer sufficient value for money or generate enough wider benefits to be worth continuing. The skills of the potential user base should therefore be an essential consideration in the design of any open data project.

Matching supply and demand

Research notes
The **Salford Data Quay** invested resources in creating **5 star linked data** before realising there was a **lack of demand** for this data. There are many different types of data users, and open data projects need to be aware of the audience. It may be beneficial to **segment data users** through providing information in different formats.

Several case studies and discussions at the open data workshop raised the problem of aligning supply of, and demand for, open data. This is an important consideration in terms of making the initial case for opening data, but also in terms of longer term sustainability.

In the initial stages of open data, it can be difficult to decide what data to release without first understanding the demand for data. Yet often external demand and interest in open data can be difficult to estimate without an existing supply, or may first have to be fostered. This can essentially lead to a 'chicken and egg' situation, whereby data users and data sources are mutually dependent. Without first creating or understanding the existing demand for open data, it can be difficult to justify investment. This challenge often creates a barrier to opening new data.

A mismatch of supply and demand can also prove challenging to the long-term sustainability of existing open data projects, even if they have high profile backers and substantial funding. This was the case for the Copenhagen City Data Exchange (CDE) – a data portal for open and closed data with a commercial focus that aimed to act as a marketplace for the exchange of public and private data. As noted in the skills section above, the CDE contained a range of complex datasets that many SMEs did not have the skills to engage with, which ultimately led to lower than anticipated usage of the site. This limited usage meant that the programme failed to reach the critical mass of datasets and data purchasers needed to make the model sustainable, illustrating the sustainability challenges of misjudging supply and demand.

Similar supply and demand challenges were encountered on a more local scale by the developers of the Salford Data Quay (SDQ). The SDQ initially invested significant time and resources in creating high-quality 5 star linked data, yet subsequently found there was no particular demand for this in the user community – 5 star linked data required more advanced technical skills that were unreasonable to expect from those regularly interacting with the data. This ultimately meant that the project had to refocus the platform on high-quality, easy-to-use datasets instead. The challenges faced by the SDQ further highlight the difficulty of aligning supply and demand. In this instance, holding overoptimistic expectations about the level of interest among local businesses and the developer community resulted in a waste of initial time and resources and required later correction.

Communication and awareness

Research notes
Teams like NYC Analytics and NOLAlytics perform **data drops that promote the availability and use of potential datasets**. This can help to increase awareness and usage. **Communication is two-way.** Open data projects have found that some data releases have led to little usage or uptake. **Alerting potential users to datasets** – including their release, and how they might be used – can be more effective when the data released has been identified through **listening to data users**.

Communication between the owners and potential users of open data is key to drive up its usability and usage, but this can be a challenge. There was recognition at the open data workshop that communication between the public and private sectors can be limited – often public sector user engagement is internal or with a set group, but does not involve speaking to external organisations about the information they may need. Private sector organisations also generally have very little engagement with the public sector over requested data. This implies that businesses may not know what data is available, or that when they do request data the ask does not always get shared. This is essentially an asymmetry of information in both directions: the public sector arguably does not know what data businesses want, and businesses are unlikely to know what data the public sector has. Without addressing this asymmetry, broader engagement with public data by private organisations may be limited.

Beyond this asymmetry of information, another key challenge relates to communicating data releases to the potential user base. Participants at the workshop noted that some open data initiatives rely solely on social media to communicate with the user base, but this can exclude some potential users as not everyone uses these social communication channels. Inadequate communication of data releases will limit open data usage, which limits its overall value.

Section 4: What does an open data project need to become a success?

This section identifies the essential components of any open data project. It provides examples of best practice from the collection of case studies and wider research, and summarises these findings into an overarching open data ambition for each of the following categories: data; data management; leadership and strategic alignment; culture; sustainability; licencing and standards; communication and awareness; and commercialisation. Examples of key actions to achieve these overarching aims are then listed for each category – which should also mitigate some of the challenges identified in the previous section. These actions are organised into three priority groups, using a MoSCoW-type model:

- **Must have:** these are essential or necessary components of opening up data, and specifically for any project seeking to open up more data for re-use. As such, they are the bare minimum required to open or release information for re-use.
- **Should have:** these are components for increasing the usability and usage of open data. They are important actions to take, and though not vital, can add significant value. These components make an open data project of good quality, long-lasting and support overall impact.
- **Could have:** these are components for creating a good open data ecosystem. These components are more about connectivity, and ensuring the wider economic and social impact of opening up data. If these components are not pursued there will be little overall impact on opening up data.

It is worth noting that these groups of actions are based on findings and examples identified through the case studies, open data workshop and wider research to date. As a result, they are indicative rather than exhaustive, and can be subjectively amended. Our approach has been to identify what is required at a minimum, what makes opening data better, and how any place can be one of the best digital city regions for open data.

Data

First and foremost, data must be available, of good quality, and usable. Given the large quantity of data that could theoretically be published, it is important to take a systematic, or curated approach to data publication, prioritising datasets according to their potential value. Several case studies illustrate possible approaches to this. For instance, the London Datastore increasingly takes a challenge-based approach to data publication, focussing on the most useful datasets that can help overcome specific urban challenges, such as inequality or net zero. This process is informed by wide consultation with relevant potential data users.

Any local data programme needs consistency in the frequency and quality of data updates. A rolling update process, as seen in the OS Open Data case study, helps give data users the confidence that they can rely on a data portal for the most up-to-date and highest quality data available. This should be automated where possible, but the process should also be regularly evaluated to make any necessary amendments based on feedback or troubleshooting.

Data leaders must also consider in what form data should best be published and take these decisions based on identified demand. Visualisations can be very useful to enhance use and understanding of data, such as key city metrics on city dashboards provided by Smart Dublin or DataDriven. Several case studies included interactive maps, allowing users to explore data on their own terms rather than being limited to the questions posed by data publishers. These maps showcase the potential value of linked data.

However, it is also important to recognise that many developers and analysts may actually prefer raw data files that they can manipulate themselves. It is important to be realistic about what data users really want. Developers may be more interested in raw data in the form of CSV files, that they can then manipulate as they see fit, rather than 5-star linked data. For example, NYC Analytics cultivates a very large user base with little visualised content.

There is a need for decisions around what data to open to be a gradual, iterative process informed by continuous dialogue with the local data community. Where a local data community does not yet exist, it may be necessary to actively support the emergence of this community, as seen in the case of Data Mill North and its work with ODI Leeds. It is vital to weigh up these considerations when deciding where to allocate resources within a local data programme.

Overall ambition: Ensure that open data is available, accessible, good quality, and usable.

Example MoSCoW actions to support this ambition:

Must do	Should do	Could do
• Source and identify datasets to publish • Create a website or other location to host the data • Create a pipeline of work to support more open data releases	• Run regular consultations with key stakeholders to identify the most useful and needed datasets • Create a complete and comprehensive list of metadata	• Create visualisations and interactive maps to enhance usability and allow users to explore data on their own terms • Publish contextual narrative about the background of datasets

Data management

Strong data management and information governance processes are essential to the success of an open data project. A range of different approaches to data management and information governance were identified across the case studies and wider research. For example, having designated data stewards and data champions is one way to encourage effective data management. Maintaining and opening a full Information Asset Register with designated responsibility for each dataset is another possibility, which helps to support information governance considerations. Embedding an information governance lead within the open data team could also help support responsible management of data.

It is also necessary to have a place to store and find the data that is easily accessible. Internal data architecture needs to be robust as a top priority prior to opening data, as outlined by several participants in the open data workshop. There are many different data management options, but popular infrastructure choices appear to be CKAN, Socrata, and DataPress and AWS. For example, one workshop participant explained how AWS allows full integration of their internal data systems and front end portal, which makes data management and publication a smooth and integrated process and minimises manual effort. This type of model is likely to keep data management costs lower, proving more sustainable in the long-term. However, it is worth noting that infrastructure decisions and choices will be reliant on what stage of the open data journey an organisation is on.

Some of the wider research conversations also drew attention to the benefits of maintaining an in-house data portal rather than using an external template or provider. This gives advantages such as oversight over data quality and updates, bespoke security standards, and greater agility.

Overall ambition: Ensure effective data management and information governance procedures are embedded within the design of open data programmes.

Example MoSCoW actions to support this ambition:

Must do	Should do	Could do
• Designate responsibility for open data to a specific team • Clearly define ownership of different datasets	• Create data steward or data champion positions to ensure that data is managed effectively • Include an information governance team member on the open data team	• Build and maintain an in-house rather than external data portal to improve data oversight and flexibility

Leadership and strategic alignment

Where senior support for open data has been high in our case studies, this has been used to the organisation's advantage and called upon as a way to push the open data agenda more widely. The Food Standards Agency is a good example of this, with a detailed data strategy focussing on innovation, transparency, and reducing bureaucratic burdens for business. The agency also has data leads embedded in senior leadership.

Direction and purpose are also vital to drive the agenda forward. Open data must be aligned to clear organisational goals and strategies, and will not work in isolation. Aligning open data with wider strategy can also prove useful in getting multiple organisations on board.

Open data is not an easy or immediate win, as seen in the Salford Data Quay case study – it depends on an active user base and an open data culture. Also, as seen in several case studies, opening data is an iterative process and the benefits are often diffuse and enjoyed by external organisations. Clear aims around what the end goal

should be, backed up by specific KPIs, will help a project to stay on track and enable the development of an open data culture.

It is also clear that a well-defined organisational structure of the teams responsible for local data programmes is critical for strategic alignment. The Trafford Data Lab case study shows the benefits of a tight-knit and highly-skilled team managing a local data programme. Dependent on the organisation, other models can also be successful, and DataDriven / NOLAlytics is organisationally much more diffuse with several teams and external Data Coordinators jointly responsible for the data programme.

Overall ambition: Gain senior support and buy-in for open data, and strategic alignment of open data with wider organisational goals and strategies.

Example MoSCoW actions to support this ambition:

Must do	Should do	Could do
• Embed the open data agenda within wider organisational goals and strategies • Build long-term senior leadership support that spans beyond changes in personnel	• Create an open data policy, setting out approach, responsibilities and data to release • Design appropriate KPIs to keep track of the direction and outcomes of open data	• Designate specific responsibility for information, going beyond the traditional role of a CIO, to include open data and information governance, to a senior leader • Embed data leads within senior leadership

Culture

Clear open data principles and frameworks can support the development of an 'open data culture'. Building an open data culture is essential to ensure that local government employees and other key stakeholders have the confidence to make the decision to open more data, and see this as an integral part of their job. For instance, NYC Analytics puts a strong emphasis on its six Open Data Values. These principles encourage learning and innovating by agencies, and make data central to the way the city is governed.

A culturally embedded open data agenda is likely to encourage data managers to publish data that is of wider value, even where there is a perceived risk of data misinterpretation. Such an agenda could help to prevent the view of curating open data platforms as simply a tick-box exercise, as observed in several case studies and wider discussions. Critical collaboration, involving close interaction between those with specialised domain knowledge and those with a strategic perspective on open data, can help facilitate this cultural change.

Getting the right open data culture is also about embedding the ultimate goal of data utilisation into every step of the process. DataDriven in New Orleans has succeeded in creating an innovative and dynamic culture through a continuous focus on use

cases for data, such as predictive modelling for health and safety violations. Data Coordinators in each city department in New Orleans help sustain this by acting as open data champions within their teams. The public is also invited to pitch analytics projects directly, encouraging openness and challenge.

Encouraging interactions between data users and data managers can also be key in terms of encouraging cultural change. For example, attending and supporting local data events is one possible way to build links with the grassroots data community. These events could also provide an opportunity to further promote data releases to the community. These types of interactions can support an open data culture in two main ways: they help to maintain an internal focus on open data that further embeds it within organisational culture; yet they also generate external interest in open data that further embeds it within the wider community and wider culture. It is important to consider cultural change from both these internal and external perspectives, as maintaining interest from the user base is key to the sustainability of an open data programme.

In addition, the process of opening up data in itself can also initiate positive internal changes within an organisation. As seen in the Food Standards Agency case study, exposing datasets to public scrutiny can act as a challenge to improve internal processes and use data better, encouraging wider cultural change.

Overall ambition: Embed the open data agenda in organisational culture, building a collaborative and enabling culture that supports greater openness.

Example MoSCoW actions to support this ambition:

Must do	Should do	Could do
• Attend and support local events to build links with the grassroots data community • Promote data releases to local and national data community	• Create a guiding set of open data principles • Build up a set of use cases to articulate the wider value of open data	• Designate a member of staff in each department to be tasked with data publication • Encourage the public to pitch analytics projects directly

Sustainability

Sustainability is about a lot more than financing. Some small data programmes with relatively little funding, like the Salford Data Quay, have been successful with supportive leadership and expertise. However, long-term funding is still an important consideration. Sustained funding can provide a sense of permanence and confidence that makes a strong statement about the importance of open data. It can also make a strong statement about the organisation, or the place, as a driver of the digital sector through open data.

A measured, iterative approach, involving regular evaluation exercises, can help ensure that mistakes do not become ‘baked in’ to an open data programme. This ensures continuous improvements, sustains buy-in from stakeholders, and can help to justify continued operational funding.

Where possible, a data programme should assess what data is being used and how, as can be seen in the case of OS Open Data. Data managers should consider why some datasets are popular and others are not, and use these insights to discover which data formats, file types, and external sharing platforms should be prioritised.

Automation of data updates should be built into any open data programme early on. This can save a lot of manual work – and therefore capacity – for data managers (see Data Mill North) and also avoid issues like the accumulation of legacy datasets that are no longer used. This automation process should involve establishing seamless links between existing internal datasets and front-end portals.

To keep an open data programme active, it is vital not to see data publication as the end goal. Improving and increasing the usage of existing data sets is equally as important as increasing the overall size of the data catalogue. An active Use Case Typology, such as that used by DataDriven, can help maintain a focus on this goal. The DataDriven Use Case Typology describes six possible opportunities for data analytics to be used in government, focusing on how existing datasets offer practical opportunities in terms of improved services and city outcomes. Alongside this, DataDriven also monitors levels of perceived interest in its data through a comprehensive Data Inventory, which further maintains the focus on driving up the usage of existing data.

Overall ambition: Achieve long-term certainty and sustainability of open data initiatives.

Example MoSCoW actions to support this ambition:

Must do	Should do	Could do
• Obtain long-term funding for open data	• Automate updates as far as possible • Review the use of published data and make relevant changes to data publication	• Create and maintain a Use Case Typology to keep track of the usage of existing datasets

Licencing and Standards

Data standards are important, especially across multiple districts and agencies. While it can be easy to set and maintain standards within an organisation, it can be much more difficult to achieve consistent standards across multiple organisations. This problem can be partly mitigated by creating a data standards, or data quality, checklist that all data owners must complete in order to publish a dataset on the portal. Making these standards open, and working with partners over time to develop and embed these standards in data collection, can have a significant impact on the usability of data.

Whilst our case studies identified that it is often overlooked and viewed as time-consuming, good quality metadata can make a dataset much more valuable. Links to data sources, contextual data, and references to other datasets improve usability, accessibility, searchability and allow data to be utilised in a more connected way. The FSA data catalogue provides an effective benchmark for metadata quality.

Drawing attention to the practical benefits of metadata, as well as employing a consistent template, may help data publishers pay more attention to this task. Guidelines such as the EDP's Metadata Quality Assurance Framework can also help to improve the consistency of metadata standards across datasets, covering the following measures: interoperability, findability, contextuality, accessibility and reusability. However, having a framework in place is not necessarily enough to ensure it is applied in practice – therefore processes should be drawn up to ensure that data publishers comply with desired standards and frameworks.

Within the UK, the Open Government Licence standard for licensing data should also help to maintain a basic standard for licensing data. While this doesn't address variation in data or metadata standards between organisations, commitment to the Open Government Licence should be an essential step in any consideration of opening public sector data.

Overall ambition: Ensure adequate data and metadata standards and licencing of open data.

Example MoSCoW actions to support this ambition:

Must do	Should do	Could do
• Create a basic data quality standards checklist • Commit to the Open Government Licence standard for licencing data in the UK	• Create an organisation or region-wide governance framework for data and metadata standards • Create a process to ensure that data publishers follow the basic data standards checklist • Develop and agree set data standards and schemas.	• Give clear consideration to the balance between portal size and its subsequent impact on standards • Create use cases to illustrate the value of standards to motivate data owners to subscribe to higher standards

Communications and awareness

A strong communication strategy is vital to the success of open data initiatives. This should focus on making data available and effectively communicating the availability of that data. It could also include plans to ensure that open data is easy to find and navigate, with filters, text searches, categorisation, and links between related pages. Where information is stored in different places – as is the case for a city region like Greater Manchester, which has data stored on data.gov.uk, local authority pages, the TfGM website, and MappingGM amongst others – a single metadata catalogue also becomes essential for navigability and ease of use. London Datastore is a good example of a datastore that effectively brings together region wide data from different sources in an easily navigable way.

It is worth investing time and resources in communication and publicity measures to ensure high discoverability and utilisation of key datasets. A good example of this is the Food Standards Agency, who have developed internal measures to ensure that

every new data release is shared and distributed on the appropriate media channels, including by e-mail, to those who have subscribed to updates. This type of communication also sends a wider message that can boost an organisation's image and reputation for being an open and digital place.

As well as communicating the availability of the open data itself, it is also important to communicate what an open data project is trying to achieve, so that potential users understand what is on offer. The case studies have outlined a wide range of effective communication approaches that raise awareness of data availability and its potential uses, such as 'data drops'. Data drops essentially involve publicising new datasets on a blog or social media channel to raise awareness of the new data and communicate its broader purpose to potential users. TfL is a good example of an organisation that makes regular 'data drops' on its Digital Blog, and recently posted a 'data drop' on the release of Electric Vehicle Rapid Charging Points data. This blog post outlined relevant details like what the rapid charge points are, how they can be used and paid for, and where the data have come from.

As part of wider awareness-raising, several data projects also hold different events to actively increase demand for high-quality datasets and find new ways of using them. 'Hackathon days' are a good example of such an event – Hackathons bring developers together, often over a day or two, to rapidly create new software from scratch using existing data. For example, during a 2014 Public Sector Hackathon, a new API for data on river levels in the UK was created. This API offers huge value to those at risk from flooding, climate scientists and others interested in flood risk management, and effectively increases usage of this existing data by improving the ease of access.

Similarly, together with ODI Leeds, Data Mill North organises 'Innovation Labs' to help solve particular data-related challenges and offer fresh perspectives on data from the wider community. These labs will be framed around a challenge, such as 'how can we make travel information accessible to older people?' and involve teams getting together to explore different options and solutions using existing data. In the case of age-friendly transport, this innovation lab led to the development of Bus Beacons – a portable device that shows users the amount of time remaining for their bus to arrive at their nearest stop that can be installed in individual homes or care homes. By making this data accessible to elderly people without the need for expensive or complicated technology, the Bus Beacon gives older people more confidence in using public transport. These types of Hackathon and Innovation Lab events can be a good way to overcome inertia and foster a dialogue between data managers and data users, and can also support the longer term sustainability of an open data programme by driving up data usage and awareness.

Overall ambition: Ensure easy discoverability of data and clear messaging around the value and purpose of open data.

Example MoSCoW actions to support this ambition:

Must do	Should do	Could do
• Create a strong communication strategy to guide data releases	• Create 'data drops' to communicate new data releases more widely	• Design and hold hackathons to increase demand for

Must do	Should do	Could do
• Develop a central repository of metadata for all datasets, where information is stored in different places	• Build a feedback function to gain insights into how users interact with the site and data	datasets and find new uses for them • Design and hold innovation days

Commercialisation

As noted previously in this report, open data programmes often have their origins in the public sector transparency agenda and have rarely had an explicit focus on commercialisation. The idea of commercialisation is therefore still relatively new terrain, especially within local government. When asked about engagement with the private sector about the data they would most like to use, most of our case studies and workshop attendees identified that they had not done this previously.

The most often cited datasets with a potential for commercialisation largely relate to transport data. There is an existing high demand for good quality, real-time data, and transport data is already systematically collected by transport providers. Developers have been highly successful in taking advantage of the commercial opportunities of TfL Open Data. However, the 'Valuable data assets' listed in Section 2 of this report highlight a wide range of different data with potential commercial value identified through this research.

As mentioned in the challenges and barriers section, it is important to consider the distributional impact of data commercialisation and ensure that benefits can flow to smaller as well as larger organisations. Data leaders should try to maintain a level playing field and make data accessible to as wide an audience as possible – for example, through creating measures to support all businesses in accessing and using complex data.

It is also important not to make assumptions about what kind of data will be perceived as valuable. Any data programme should keep in touch with the business community to keep track of data demand as it emerges and develops. These conversations should then feed into a prioritisation matrix for new data requests.

Overall ambition: Encourage commercial engagement with public sector open data, actively seeking to identify and address demand.

Example MoSCoW actions to support this ambition:

Must do	Should do	Could do
• Hold conversations with the business community to establish the needs of the potential user base • Create a prioritisation matrix for new data requests	• Consider the equalities impact of opening data	• Identify and create measures to support SMEs in accessing and using data

Section 5: Where next for Greater Manchester?

We want Greater Manchester to open data that has the biggest potential value for our region, and make the best use of that data. To become a digital city region with an evolving and world class digital ecosystem, there are a series of priorities that Greater Manchester must have, should have, and could have. The previous section of the report outlined an indicative list of must haves, should haves and could haves for open data in a broad and non-specific context. Many of these identified actions apply to Greater Manchester too, so there is some overlap with the points mentioned here. However, this section draws out the must haves, should haves and could haves for Greater Manchester specifically, based on our understanding of where the region is currently at in relation to open data.

The main deliverables and actions for Greater Manchester at this stage in the research are outlined below. It is important to emphasise that these priorities and actions have emerged from the case studies, open data workshop and wider research to date, but do not yet indicate the exact work that needs to be done. Rather, they indicate current perceptions about what must, should and could be done from an internal public sector perspective, and are indicative at this stage.

The next workstream in the Local Data Review is an external consultation on open data, which will draw out additional ideas and priorities according to the perspectives of a range of external business, academic and voluntary organisations. Once this consultation has been conducted, the priorities listed below will be added to and amended to create a final prioritised list of actions and deliverables for Greater Manchester, reflecting both internal and external perspectives together.

Must have

The list below shows the basic deliverables that Greater Manchester must have in order to open data. Most of these points are relatively simple and relate primarily to stage 1 (data setup) and stage 2 (data quality) of our general open data progression (see section 2 for further details). Some of these basic steps are already in place, while some steps will need further action to achieve.

- Build on the success of open data projects in Greater Manchester by prioritising and opening up more, simple datasets that users want and need. This can help to grow demand, and stimulate the view of Greater Manchester as an open, digital place.
- Develop and agree upon a set of basic standards for the highest priority datasets. This should ensure data is consistent and comparable across Greater Manchester, adding value to the datasets and supporting their wider re-use.
- Organise a programme of work to support more open data releases and to effectively identify and prioritise new data releases.
- Create an easily accessible place to share open data for each public sector organisation in Greater Manchester, building on the good practice and learning of organisations like Trafford Data Lab, Salford City Council, Stockport Council, Wigan Council and others.
- Establish a basic central repository of metadata that describes all datasets held by the public sector in Greater Manchester. This aim would be long-term

and could be based on information asset registers. This action would help to identify to potential users datasets that they might not have previously been aware.

- Ensure open data is provided to at least 3 star level and in more than one format. This means all data should be provided in a non-proprietary open format that does not require any particular software package – for example, using CSV format instead of Microsoft Excel's XLS or XLSX formats. This can enhance usability, and thereby usage, of new open data.
- Establish an inclusive governance framework for open data across Greater Manchester. This is already in place at an organisational level across many Greater Manchester organisations, but there is currently no GM-wide governance framework specifically with an open data remit.
- Lay out an open data roadmap for the Greater Manchester public sector, to guide organisations through the progression of simply opening data, to focusing on increasing data usage and value.
- Articulate clear, consistent messaging around the value and purpose of open data, linking this to strategic aims from the Greater Manchester Digital Blueprint, Local Industrial Strategy, and Greater Manchester Strategy. This should be supported by a strong communication strategy to guide data releases.

Should have

Beyond the core essentials for opening data outlined above, the list below outlines some of the things that Greater Manchester should have in place to make sure that data is not simply made open, but adds value by being useable and used. These elements should shift the focus towards making open data higher quality, more valuable, and long-lasting.

- Create a pan-Greater Manchester portfolio for data and information, covering data use, information governance, research and analysis. To ensure sustainability of this approach, efforts will need to be made to embed open data priorities within wider organisational culture and strategies across the region.
- Create a standards board for Greater Manchester, paying consideration to the trade-off between capacity and standards. The larger the number of organisations involved in sharing data, the harder it can be to maintain high and consistent standards.
- Create a Greater Manchester-wide datastore. This would need strategic buy-in from leaders across the region to ensure its longer-term sustainability, given the previous experiences of DataGM and GMDSP. It would also need thorough consideration of the the choice of underlying digital infrastructure, such as CKAN, Socrata and AWS, and should have automation built in as much as possible.
- Establish a Greater Manchester-wide process for opening data that includes information governance considerations.
- Expand the level of detail in the central metadata repository, representing an awareness that links to data sources, contextual data, and references to other datasets improve usability, accessibility, searchability and allow data to be utilised in a more connected way. Wigan Council Open Data is a good

example of a data portal that includes comprehensive metadata – this level of detail would be beneficial to create on a Greater Manchester scale.

- Have designated data stewards or data champions and embed an IG team member in the open data team in organisations across the region to help ensure that data is managed effectively.
- Emphasise the curation, rather than simply the opening, of datasets in order to maintain the value of data over time.
- Agree an information strategy across the region, which aligns to wider regional strategies and has senior leadership buy-in. This should outline a clear direction and purpose for the end goal of open data and should be supported with key performance indicators. These indicators should help the project stay on track and help track progress and outcomes over time.
- Review the use of published data and design evaluation mechanisms to understand how data is being used. Use the outcomes of these evaluations and reviews to make relevant changes to the prioritisation and publication of data.
- Ensure each organisation has set their own open data policy, including the approach, responsibilities and prioritised data to release.
- Widely promote open data through approaches such as ‘data drops’, or by sharing open data releases and use cases on social media.

Could have

There are features beyond those listed above that Greater Manchester could have in order to create a better open data ecosystem and culture of openness in Greater Manchester. On the whole, these steps relate more to stage 4 (data usage) of our general open data progression model, and attempt to drive up the usage of open data across the region. These are the wider steps that form part of a longer-term strategic approach to make Greater Manchester known as a digital city region with a world class and evolving digital ecosystem.

- Build visualisation and open analysis features into the datastore to allow users to explore the data on their own terms, rather than on a predetermined set of options. This could be supported by contextual narratives, like the Story Maps feature in Wigan Council Open Data, which may improve the usability of data and make it accessible to a wider audience.
- Actively collaborate with the Local Enterprise Partnership, Chamber of Commerce and other private sector organisations to discuss ideas and preferences for potential new datasets, and ascertain levels of demand from the business community.
- Create an agreed Greater Manchester-wide open data policy, setting out the approach, responsibilities and prioritised data to release. This should more firmly embed an open data culture across the region.
- Engage with the open data community through approach such as hackathons or innovation days. This should increase demand for datasets and find new ways to use existing data, and offer fresh data perspectives from the wider community.
- Encourage developers to engage with data and discover new ways of using it. For example, Trafford Data Lab has a ‘showing their working’ function to systematically explain some important techniques used by the team.

- Migrate to an in-house data portal over time, where the content and infrastructure is controlled by the public sector. Some wider research discussions have noted that bringing a data portal in-house allows it to be aligned with the internal data systems. This can have significant benefits in relation to data oversight, flexibility and information governance.
- Open a Greater Manchester Information Asset Register with a full list of all datasets held in the region, both open and closed, along with designated responsibility for each dataset.
- Work towards providing data to 5 star level open data as skills and awareness within and outside the public sector matures. This means data are in a non-proprietary open format, as noted previously at 3 star level, but also have a URI and can be shared on the Web, and are linked to other data to provide context.
- Embed data leads within senior leadership to help with general oversight of data and effectively ensure that data meets the expected standards and quality.
- Consider segregating the front end data portal into different sections to cater to multiple audiences, with greater visualisation and open analysis to suit non-specialist users, and aggregated raw data for developers that can be accessed using APIs.

Next steps for the Local Data Review

As noted previously, the deliverables outlined above are indicative at this stage. They are also largely based on an internal public sector perspective, which is only one piece of the puzzle. It is vitally important that we also begin to understand the challenges data users face in accessing and using our data.

Our next steps will therefore be to develop this understanding of the external use of public sector data, through running an external business engagement exercise. This work, conducted by Open Data Manchester, will seek to gain an understanding of the external perspective of public sector open data: the challenges in accessing and using open public sector; and the opportunities to make accessibility and use of our data easier. While this work will predominantly focus on the business and commercialisation uses of public sector data, the aim will also be to understand the innovative academic and third sector challenges and uses as well.

The list of priorities and actions for Greater Manchester outlined above in Section 5 will therefore be amended and added to, following this exercise. Ultimately, we will endeavour to develop a prioritised list of deliverables for Greater Manchester that should help us to open more data; ensure that our data are useable and used; and create one of the best open data ecosystems in Europe.

Appendix 1: Full List of case studies and their relevance to the Local Data Review

Case Study	Scope	Relevance to Local Data Review
DataGM	Local	Earliest attempt to create a region wide simple datastore for Greater Manchester. Its closure in 2018 provides key lessons for future projects in the region.
Greater Manchester Data Synchronisation Project	Local	Aimed to link datasets in Greater Manchestter from a range of different sources and publish these openly, with the intention of improving local administration and driving innovation. Its closure in 2019 also provides key lessons for future regional projects.
Trafford Data Lab	Local	Local data portal that has a strong focus on the reproducibility of data and reuse of data by developers, which includes devoting time to 'showing their working' for others to use and learn from.
Wigan Open Data	Local	Local data portal that was a staff-led initiative, underpinned by a comprehensive Data Visualisation Strategy. Pioneers for GIS and open data in the region.
Salford Data Quay	Local	Local project that has built and maintained a well-functioning datastore with modest seed money. Built in automation to ensure that data remains usable and high quality, without requiring significant manual input.
Stockport Open Data	Local	Local council that is currently focusing on the internal robustness of data systems as a key first step in opening data. Offers valuable insight and lessons for a future data portal from the previous attempt at the MyStockport open data platform (now discontinued).
Data Mill North	Local and national - some local data from Stockport included	Large datastore run by Leeds City Council. Focuses on an open and collaborative approach to stimulate creative uses of data.
MappingGM	Local	Currently the home of geospatial data in Greater Manchester, offering a simplified user experience and clear visualisation of data.

Case Study	Scope	Relevance to Local Data Review
Open Data Manchester	Local, national and international – ODM works with multiple partners around the country and internationally	Community Interest Company that works with a range of councils, local communities, universities and private companies to support responsible and intelligent data use. Developed an interesting Declaration for Responsible and Intelligent Data Practice.
Insight Cheshire East	National – a local authority programme outside of GM	Contains examples of potential innovative features for a data portal – such as Story Maps of a 'Create a Webmap' function. Illustrates good practice in terms of information governance and management through the 'Information Asset Register'.
TfGM Open Data	Local	Greater Manchester's transport body. Has a strategic focus on using open data to hlep customers make informed travel choices, and on the wider regional economic benefits of open transport data.
London Datastore	National	Successful repository for data sharing between numerous local authorities, emergency services and third sector organisations. It has also built a strong external reputation of the London city region being an open and innovative place.
TfL Open Data	National	One of the earliest organisations in the UK to open its data. Looks for new ways to use its data to improve quality of life and create economic value, and has a strategic focus on commercialisation by private sector organisations.
Food Standards Agency Open Data	National	Open data is influenced by its strong Data Strategy, which embeds open data within senior leadership and organisational culture. It also has a robust open data engagement strategy.
Ordnance Survey Open Data	National	Provides a range of geospatial data about Greate Britain. Focuses on openness within its business model to drive innovation and growth.
Data.gov.uk	National	National datastore that was initially driven by the government transparency agenda, but now incorporates economic objectives through its strategic drive for public data to be used in innovative was.

Case Study	Scope	Relevance to Local Data Review
NYC Analytics	International	One of the largest and most well-used open data platforms in the world. Arguably a beacon of best practice for open data portal design and governance.
DataDriven and NOLAlytics	International	Aims to find data-driven solutions for addressing city problems and focuses on public engagement.
US Government Open Data	International	Large catalogue of US government datasets, with high data uptake by developers. Offers some lessons around data standards and quality issues due to its large scale.
Australia and New Zealand Infrastructure Pipeline	International	Provides a forward view of public infrastructure activity across Australia and New Zealand with the intention of drawing private sector investment into the region.
National Map (Australia)	International	Intends to support commercial and community innovation through the opening of public data, with a strong focus on infrastructure and investment spatial data. It includes an innovative 'Investor Map' to encourage investment across the country.
European Data Portal	International	Created to harvest the metadata of Public Sector Information across European countries. Covers a huge geographical area and range of organisations, illustrating the trade-offs between quantity and quality.
European Union Open Data Portal	International	Provides a single point of access for European Union data, serving the twin goals of boosting economic development and improving transparency of EU institutions.
Dublinked	International	Driven by a strong digitisation agenda and focuses on innovative uses of data to improve services. It also coordinates the four local authorities in Dublin to pool resources and solve problems collectively.
Copenhagen Data Exchange	International	Created a market for selling and purchasing data, acting as a matchmaker between data suppliers and consumers. Its discontinuation in 2018 provides valuable lessons for the future, particularly around matching supply and demand.