GREATER MANCHESTER

DOING THINGS DIFFERENTLY FOR OUR ECONOMY

Local Data Review

Final Research Summary

20th April 2021

Contents

Executive summary

Greater Manchester aims to be one of Europe's top digital city regions. The world's best digital city regions use their data effectively to support growth and innovation. Greater Manchester needs to do the same.

Our digital city region ambitions are laid out in the Greater Manchester Digital Blueprint and Greater Manchester Local Industrial Strategy. But the foundations of these were laid in the Independent Prosperity Review, where a panel of eminent economists highlighted some of the main challenges we face in Greater Manchester – and how we can support the digital and digitally-enabled sectors in Greater Manchester to grow and innovate.

The Local Industrial Strategy set out the role of open data in supporting Greater Manchester. We want to open data that is of value to Greater Manchester, drives productivity growth and innovation. To meet these ambitions, we must first build an understanding of the open data ecosystem in Greater Manchester. We must also learn from the successes, challenges and experiences of others when opening data – both within and beyond Greater Manchester. This Local Data Review is an attempt to build this understanding and identify what we must do to create a better open data ecosystem.

Our Local Data Review looks at the challenges we face in opening more data for re-use, and the things we need to do to overcome those challenges.

The review explores these challenges from two key perspectives: the internal perspective of public sector organisations, gained through a series of case studies and an open data workshop; and the external perspectives of businesses, academia and the third sector, gained through an external consultation exercise. This report concludes our Local Data Review by bringing these two perspectives together and drawing out the key lessons learned.

We found that opening data requires an iterative approach, gradually opening and improving the quality of data. Some organisations are further along this path than others.

Early-stage open data projects tend to focus more on increasing the amount of data they make open. More advanced projects pay greater attention to improving the quality and usage of their data. This suggests a general path of open data progression over time. Organisations that focus too heavily on a later stage alone have often found it difficult to progress. By contrast, the most successful open data programmes have paid attention to all stages in a gradual and iterative approach that tackles different elements and challenges along the way.

The challenges faced in releasing data are varied and often linked.

The main challenge is often maintaining long-term sustainability. Open data programmes can become unsustainable for a vast range of reasons – for example, due to lack of leadership buy-in or cultural resistance to the open data agenda. Challenges around the selection, management, and upkeep of the data itself are also common internal barriers. These data issues link to further commonplace challenges, with the supply of and demand for open data closely intertwined.

To tackle these challenges, we must focus on understanding how data might be used and prioritise the release of data that has the greatest impact.

Maintaining continuous engagement between data users[1] and data controllers will be key to building this understanding. This review has initiated dialogue with local businesses and other data users, which has provided an overview of how some data is currently used in Greater Manchester. However, this is the first step in a much longer process – these conversations must be continued to prioritise the release of data that has the greatest impact.

We identified several thematic areas where data releases would add significant value by supporting innovation and productivity growth.

These broad data thematic areas are transport; business rates; socioeconomics and demographics; planning and housing; movement and activity; asset location; COVID-19; infrastructure; environment; health; education; and safety and security.

[1] See Appendix 1 - Note on terms and definitions

But simply making data open does not mean anyone will use it. We explored what prevents businesses, academia and the third sector from using public sector data.

Challenges that prevent these organisations from using public sector data often relate to the data itself, and the real or perceived lack of quality, consistency, timeliness, availability, and stability of that data. Further challenges around accessing data, primarily due to lack skills in both producing and using data, can also limit the re-use of public sector data. These issues need to be tackled in tandem.

We also looked at what we could do to increase data access and usage.

Maintaining continuous engagement with data users will be key to increasing data access and usage. We found that local data was perceived to carry greater levels of risk than national data sources for data users – such as poor quality or incomplete data requiring significant capacity to 'wrangle' into a usable format; or risks around products or services that are built on data that is not renewed, or taken offline. To increase local data re-use, we must remove these risks by prioritising the release of high quality, standardised, and reliable data, good documentation, and updates on availability.

These findings indicate that we must 'level up' Greater Manchester by focusing on the basic 'open data foundations'.

Our Local Data Review has outlined the main open data challenges for the city region and identified 12 essential deliverables to overcome these challenges. These deliverables include developing a roadmap to 'level up' organisations – taking them through the progression from the initial release of open data, to increasing the quantity, quality, and usage of that data.

While these deliverables may seem ambitious, they should be approached iteratively and rolled out gradually, following a set action plan. They are listed below in the order they should roughly be carried out, as they largely build on one another. Our recommendations are:

1. Open more datasets, from more organisations, that users want and need.
2. Put in place a programme of work to identify, prioritise and support new data releases.
3. Create an easily accessible place to share open data for each public sector organisation in Greater Manchester.
4. Develop and agree upon a set of basic data standards and documentation for the highest priority datasets.
5. Establish a central repository that lists and describes all datasets held by the public sector in Greater Manchester.
6. Agree to open data in accessible, open formats.
7. Establish an inclusive governance framework for open data across Greater Manchester.
8. Create an effective communication strategy that articulates the value and purpose of open data.
9. Listen to and engage with data users, starting now.
10. Make the case for open data releases that will support business growth and innovation in Greater Manchester.
11. Ensure open data is embedded in all policy areas in Greater Manchester.
12. Lay out an open data roadmap - an action plan - for the Greater Manchester public sector.

Yet opening data is not an end in itself.

The 12 essential deliverables are the first step in a much wider process. We have drawn up a series of ‘should have’ and ‘could have’ deliverables to support and expand upon this initial set[2]. Our next steps will be to create an action plan that sets out how to achieve the essential deliverables and open up more priority datasets. This will be further supported by wider considerations of data quality and the importance of public sector transparency.

Our action plan will not only help to open more data across the region but will also support the creation of a better information ecosystem.

This should drive us closer towards our goal of becoming a leading digital city region, as set out in our Local Industrial Strategy and Digital Blueprint.

[2] Outlined in Annex A

Background and context

Introduction

Greater Manchester aims to be one of Europe's top digital city regions. Evidence suggests that the leading digital city regions make the most effective use of their data. To become a leading digital city region, Greater Manchester must do the same. Making the best use of data does not merely mean organisations using their own data well. It also means ensuring organisations can access and use data that they do not collect or store. Data can add value to processes, products, and services when it is used. Opening 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.

Strategic context

The aim for Greater Manchester to be a top digital city region has its origins in the Independent Prosperity Review, the Local Industrial Strategy and the Digital Blueprint. In March 2019, the Independent Prosperity Review set out the current and future state of Greater Manchester's economy. Led by a panel of eminent economists, the review highlighted some of Greater Manchester's existing strengths, particularly around health innovation, manufacturing, digital and creative industries, and professional services. However, it also concluded that more needed to be done to reduce social inequality, drive up wage growth, and boost productivity in the region.

The Local Industrial Strategy (LIS), published in June 2019, responded to many of the issues highlighted in the Independent Prosperity Review and sets out plans to drive Greater Manchester's economy forward. The LIS recognises that digital technology creates significant opportunities to raise productivity and competitiveness in the city-region. It also highlights the importance of local data to stimulate open innovation and enable productivity growth. This leads the LIS to suggest that data that can be made available should be made available, unless it is prohibitively expensive or not appropriate to do so.

These ideas are also echoed throughout the Digital Blueprint, which aims to make non-personal data open by default where it is of value to the people and businesses of Greater Manchester. Opening up local data in this way will be key to achieving the Digital Blueprint's broader ambition of making Greater Manchester a top five European Digital city-region.

The Local Data Review

The Local Data Review sits within this wider strategic context and is a direct action of the Local Industrial Strategy. However, it also directly addresses priorities and challenges set out within the Independent Prosperity Review and Digital Blueprint.

The Local Data Review aims to understand the challenges of making more data open and recommend actions to address these challenges. The public sector, as a significant data producer, can help to develop a strong open data ecosystem in Greater Manchester. Doing so would create new commercial opportunities, stimulate innovation, and raise productivity – supporting wider progress towards a more prosperous city-region.

What is this report?

This document is the summary of the research and analysis pulled together for the Local Data Review. It sets out some of the most important challenges from two perspectives. The first is that of the public sector, and the internal barriers the public sector faces in making data open. This internal perspective was gained through a detailed case study analysis and workshop. The second perspective is that of businesses, academia and the third sector, and the barriers they face in accessing public sector data. This perspective was gained through an external consultation exercise with a survey and a series of focus groups.

The recommendations in this summary highlight what must be done to address open data challenges faced both internally and externally. By addressing these challenges, the public sector can increase the usability, use, and value of our major public data assets.

Case study analysis

Methodology

Our initial work involved understanding the challenges faced by public sector organisations trying to make data open. This work included:

- A literature review conducted by IDOX.
- A range of 25 case studies from Greater Manchester, the UK and internationally, which were informed by the literature review and local contacts.
- A workshop presenting the findings of the case studies, which also explored the actions that Greater Manchester might take to make data more open.

While our work focused on identifying lessons learned from these case studies, they also provided a range of contextual information that informed the work in more detail.

Evolution of data projects

One key finding was the general progression of open data projects. Four simple stages were identified[3]:

1. **Beginning to open data** – this stage simply requires some information to be released openly.
2. **Expanding the amount of open data** – projects that had established the necessary elements to open data could then expand the range of data made open and explore new datasets.
3. **Improving the quality of open data** – once larger numbers of datasets were released, the focus of the work sometimes shifted towards improving the quality of data released, to improve its value and usefulness.

[3] See Appendix 2 for further details of these four stages.

4. **Increasing data usage** – projects relied on excellent quality open data to focus on increasing usage, engaging more with users, and supporting events like hackathons.

Some of the case studies explored were further along this progression route than others, while some incorporated all aspects. The general progression along this path was dependent on each step being taken, rather than jumping straight to a later stage. Projects that tried to focus on a later stage were often incredibly innovative and considered ahead of their time, but without a specific focus on earlier stages and the building up of use cases, they found it difficult to progress.

This suggested that simply opening data, while possible, is not sustainable in the long-term without a more iterative, staged approach. This requires proper evaluation of what has gone before and developing use cases for open data. The focus of this work – releasing data for wider reuse to support innovation and productivity growth – could function as a good approach for developing use cases with private, academic and third sector organisations[4].

Challenges and barriers

Challenges and barriers around opening data fall into two main categories: perceived and real. We noted several recurring examples of perceived barriers, such as fears surrounding potential misinterpretation and misuse of data or concerns about information governance being a barrier. While these perceived barriers rarely transpired into practical issues, we also noted numerous genuine practical challenges to opening data.

Long-term sustainability

Long-term sustainability is often the main challenge for an open data initiative. Open data projects can become unsustainable for a vast range of reasons, from a lack of long-term financing, to insufficient internal capacity and resources for data

[4] See Appendix 3 for a timeline of selected local and national open data initiatives from 2010 to 2020, and further narrative around how some projects have approached these distinct stages.

maintenance and upkeep, to limited demand and engagement from the data community.

Sustainability challenges often correlate to the Gartner Hype Cycle, which shows how enthusiasm and momentum for new technology often changes over time. If initial expectations around open data are too high, momentum can sharply decline when reality kicks in. This dip in momentum is commonplace within open data initiatives and for the introduction of new technology more widely. However, this loss of momentum can prove unsustainable if efforts are not made to rebuild and sustain further enthusiasm for open data[5].

Leadership

Leadership challenges can also be a barrier to open data. If long-term buy-in for open data is not gained from senior leadership, it can be hard to fully embed a culture of open data within an organisation. Where open data is not embedded within organisational culture, it can be challenging to drive the open data agenda forwards and maintain enthusiasm and momentum.

Additionally, the link between leadership and officers is also important. Trust in leadership and continual engagement can ensure that individuals collecting data understand and support the decisions made by senior leaders around sharing it.

Culture

Open data culture can be a barrier if it is framed solely in relation to the transparency agenda or is too focused on internal benefits. Such a narrow vision misses the broader purpose and value of open data. This can limit the scope for drawing up a valid open data business case based on wider social or economic benefit.

Data choice

Challenges surrounding the choice and prioritisation of which datasets to make open can also become barriers. Initial selection of data to make open requires detailed scoping and internal groundwork first. Rushing or skipping this stage can lead to

[5] See Appendix 4 for a graph of the Gartner Hype Cycle and further narrative around this open data challenge.

unnecessary duplication of datasets between national and local databases, or to the release of datasets that are not needed by the data community.

Communication and awareness

Communication between the owners and potential users of open data should be an essential part of any open data initiative, but this is often limited. This leads to an asymmetry of information, whereby the public sector often does not know what data businesses want, and businesses often do not know what data the public sector has. Failure to address this asymmetry can limit the use of open data.

Matching supply and demand

It can be difficult to decide what data to release without understanding demand, yet demand can be difficult to estimate without an existing supply. This is a ‘chicken and egg’ situation whereby data users and data sources can be mutually dependent. Overcoming this challenge can be difficult but building and maintaining demand is essential for long-term sustainability, productivity growth and innovation.

Skills

Lack of skills can be a further barrier to opening data, both internally and externally. Some public sector organisations struggle to build and maintain the skills needed to ensure data is accurate, complete, and up to date. The user community can also lack the technical skills needed to fully engage with open data, particularly if datasets are large or complex.

Mitigations to these challenges and barriers

The challenges outlined here can prevent the successful opening of data or result in long-term unsustainability of open data projects. However, there are many ways to overcome these barriers. Our full case study analysis suggests a range of actions and recommendations that can and should be taken to mitigate these key challenges.

Valuable data assets

It is difficult to learn the true value of specific data assets. Like infrastructure investment, open data can be an enabler to business growth and innovation. A range of research has tried to quantify the value – financial, economic, and social – of open data. For example, the EU estimated the market value of open data to be €184.45bn per year in 2019. Similarly, an evaluation of TfL's open data noted it supported a digital sector and helps end users to make better decisions about transport. However, this research is based on the release of data rather than a pre-release assessment.

Existing research shows that the benefits of open data go to the wider economy, rather than as significant direct financial benefits to open data producers. It is therefore important to understand the potential usage of open data, rather than its direct value. This is an essential consideration when choosing which datasets to make open.

Yet even this approach has difficulties. Our case studies showed that usage was not directly related to value – the number of data downloads does not inform *how* datasets are used, or how often, by the user. This suggests that to understand the use and value of data, there must be close engagement with users. Through doing this, we can understand how people and businesses use data, what they use it for, the challenges they face, and the demand for new datasets.

We have tried to reflect this approach in our business engagement workstream, and it is important that this engagement with users continues. By doing so, we can support a more curated approach to open data – focusing efforts on opening data that is of value to Greater Manchester.

Both the business engagement work, and the initial case study analysis, highlighted several themes where specific data releases would be valuable. These include:

- **Asset locations:** a basic set of location information on things such as street furniture, or land and property assets held by the public sector.
- **Business rates:** local authorities release this data, but not in a standardised way. It was suggested that the standardisation and release of this data would be beneficial.

- **COVID-19:** this theme includes both COVID-19 health data and relevant economic and social information, such as datasets on economic impacts, uptake of loans, and job retention scheme data.
- **Education:** this includes information such as school attendance, childhood development, OFSTED results, Early Years ASQ scores, or the destination of students post 18.
- **Environmental:** this includes a broad range of information around themes like waste, green space, air quality, weather, and flood data.
- **Health:** while health data tends to be sensitive information and subject to strict information governance procedures, there are anonymisation techniques that could enable more health information to be opened.
- **Infrastructure:** partly linked to asset locations, this theme includes information such as infrastructure locations, investment, and opportunities for development.
- **Movement and activity:** separate from transport infrastructure, this theme focuses on flows of people, use of transport systems (e.g. ticketing data), and activity monitoring, such as footfall and emissions data.
- **Planning:** information around planning activities, including future developments, and planning applications would benefit from being standardised and more openly released.
- **Safety and security:** information like police crime data and the crime survey data can have a wide range of uses.
- **Socioeconomics and demographics:** while datasets are available at a national level, subsets of national datasets tailored to local areas, would be beneficial to release.
- **Transport:** this includes real-time information on locations, as well as transport timetables, stations and stops etc.

Conclusions

The case study analysis and open data workshop allowed us to draw up a series of actions that Greater Manchester could take to open more data and build a better information ecosystem across the region. Our subsequent business engagement exercise identified further actions to add to this list. The full list of these deliverables, arranged in three priority groups, is provided in Annex A. The highest priority actions are also outlined later in this report, below the subsequent section on our business engagement exercise.

An overarching message running throughout this analysis, and underlining all the identified deliverables, is that the approach to opening more data must be gradual and iterative to be successful and sustainable. Where projects have been too ambitious too soon or did not gather perspectives from the user community before launching open data, they have usually stalled. The most successful case studies have taken a more moderate and gradual approach to opening data, with constant engagement with users. This provides space to make iterative amendments based on feedback received from both data users and data producers, and to respond to challenges that arise along the way.

A second key, and related, message is that businesses and other data users should be involved in the decisions surrounding what data should be made open. The value of data comes in its use, not in its storage. It is therefore essential that data owners and data controllers are considering the needs and preferences of those who use and interact with the data. As with the key message above, this understanding needs to be continually updated through regular dialogue and engagement with the business community. Our business engagement exercise, outlined below, was the first step in this process. This exercise has opened the door to continue future dialogue with the business, academic and third sectors.

Business engagement

Methodology

The business engagement exercise was carried out by Open Data Manchester on behalf of the GMCA. The focus of this consultation was to understand open data usage, challenges, and requirements within the business, academic and third sectors. The consultation was split into an online survey with six follow-up sectoral focus groups. Participants were acquired through social media, direct email, and the Open Data Manchester email newsletter.

Understanding of open data

There can often be confusion around whether open data includes commercial and personal data, or whether “big data” is the same as “open data”. Clarifying what open data is, and the benefits it provides, leads to more informed conversations about how to use open data.

Participants largely understood open data to be a public good – data that is available and easily accessible to all with minimal barriers. In this respect, it meant it was not restricted to information held or collated by the public sector. Emphasis was also put on the importance of data being machine-readable, supported by good documentation and comprehensive metadata. Understanding what organisations perceive as open data can help to define the approach of future open data initiatives.

The ODI Data Spectrum, outlined in Appendix 4, is a useful approach for further clarifying the meaning of open data and outlining the full range of data from open to closed.

Current open data use

The organisations taking part in the consultation used a wide variety of datasets for a wide variety of purposes. Broadly speaking, the datasets identified by participants fall across the categories of socio-economic, environmental, health, mobility, geospatial, education, safety and security, regulatory, democracy, crowdsourced, and other data.

Open data needs and uses vary somewhat by sector. Open Data Manchester therefore separated business, academic and third sector participants into different focus groups. This helped to build an understanding of sectoral-specific open data uses, needs, and challenges.

Business sector

Discussions with business respondents highlighted some of the many ways that open data is currently used within Greater Manchester's business community – from analysing the spatial impacts of climate change, to supporting infrastructure planning applications, to developing a wellbeing index.

Importantly, business participants highlighted that open data creates a low-risk environment for ideas to be explored. An example was given of a hotel booking site that wanted to use open data to explore if weather impacted booking behaviour. While this idea did not come to fruition, the availability of open weather data through an API allowed the concept to be explored. These kinds of opportunities highlight the value of open data as a tool to drive innovation and productivity growth in Greater Manchester.

Conversations with the business sector also drew out key considerations surrounding the geography of open data. Being open means that data is not constrained to a specific geographical territory. This allows applications and services developed in one area to be scaled up or applied in other locations, if data coverage and structure is consistent. As an example, transit schedule data released by TfGM was built into a Berlin-based app, Mapnificent, when the feed was made available in 2012. Recognising that local data has the potential to be used across a wider geographical base could help to showcase Greater Manchester externally as a leading digital city region.

As a further geographical consideration, while most business respondents operated out of Greater Manchester, the majority saw their market as being UK-wide. Organisational coverage was broadly reflected in the type of data accessed, and where it is sourced from. Organisations with a local coverage often sourced their data from local data portals, or from a mix of local and national sources. However, participants generally preferred national data portals when they provided country-wide services and preferred to access national datasets that could later be cut down

to a local geography. It is important to recognise the value in existing national data sources, and to ensure that local data releases do not simply duplicate national releases that are already fit for purpose. Rather, we need to prioritise data releases that have unique and specific value to our local businesses and data users.

Academic sector

Academic uses of open data identified in this consultation largely centred around research methods and training, and the use of open data within research itself.

As with the business sector, geography mattered within the academic responses. When open data is used as a tool to develop research methods and teach data skills to students, local data was generally preferred to national data. This is due to the added context that comes with living in the area, which allows for better understanding of the data. While this point was only discussed within an academic context, it is likely that local businesses would also benefit from added contextual understanding if they were to have greater access to local data.

The other key point brought out by respondents in this group was the importance of data, both open and restricted, for healthcare research. While health related data usually follows strict governance procedures, it was felt that there is scope to anonymise and open far more health data than is currently opened, or to develop a research ‘sandbox’ for health data. These ideas are outside the scope of our Local Data Review. However, we will take these suggestions forward in our wider work on building a better information ecosystem in Greater Manchester.

Third sector

The uses of open data identified by voluntary, community and social enterprise (VCSE) sector respondents primarily related to identifying need and targeting resources. Open data was also used by these respondents to create an evidence base for programme development and funding.

Most of the VCSEs spoken to in this research relied on national datasets more than local datasets, largely because national datasets offer greater comparability. The preference for national sources may also partly be due to a lack of skills and time available to search for local data within VCSEs, or due to any of the challenges outlined in the section below.

The importance of standards also emerged as a key theme within the VCSE group, who discussed the OpenActive data standard for sharing sport and fitness opportunity data. The OpenActive standard makes it easier for residents and businesses to book sports and leisure opportunities. This illustrates how efforts to build strong open data foundations and consistent standards can create value downstream, leading to practical benefits for the end user.

Challenges faced when accessing and using public sector data

The challenges faced by open data users broadly fit into two categories. The first set of challenges relate to the data itself, and the quality, consistency, and stability of that data. The second set of challenges relate to difficulties in accessing the data that is wanted or needed, primarily due to lack of timely availability or due to a lack of skills.

Quality

Concerns around the quality of open datasets are common. Often, open data appears to be missing crucial elements or does not conform to a particular standard. Lack of access to good quality open data hinders innovation and experimentation, and ultimately creates distrust of the data that is being released.

Consistency

There is currently little data consistency across the ten Greater Manchester boroughs. This limits the potential for coherent and consistent services and insights across the region.

Dependencies

As data is made available and used it creates dependencies. These dependencies can be problematic if datasets are moved or no longer shared or updated. Where this happens, applications and services built using that data can break. Without feeling confident in the stability and reliability of data releases, organisations may be reluctant to invest time in creating new and innovative uses for open data.

Timeliness

Lack of timely access to data can be a key challenge, particularly in relation to census and other demographic data. Often data is compiled at a local level and submitted to a national body for inclusion in a national dataset, such as STATS19 transport data showing road traffic collisions. Allowing organisations to have earlier sight of the local data prior to the national release could enable them to act and make decisions sooner, but this early local sharing rarely happens at present.

Skills

Lack of skills was highlighted in our case study analysis as a potential barrier to the re-use of open data. In this context, skills are assumed to equate to the capability and capacity to use open data, including knowledge, skills and attitudes. This point was raised again in the business engagement exercise, where some participants noted a willingness to use more data but lacked the capacity to do so.

Key recommendations and conclusions

This work has highlighted several recommendations that have contributed to our final list of actions and deliverables for Greater Manchester. A key recommendation emerging across this consultation is the importance of having better and more continuous engagement with data users.

Maintaining continuous engagement should build a better reciprocal understanding between data owners and data users: data users will become more aware of what open data is held within the public sector, while the public sector will become more aware of what data users want and need. These dialogues will foster gradual change towards a more inclusive open data culture.

Alongside the need for greater engagement, recommendations emerged around open data strategy. An open data strategy should be linked to other strategies across Greater Manchester, such as the emerging Information Strategy and strategies around digital inclusion. Assigning high strategic priority to open data would signal further commitment towards an open data culture, which should eventually see open data infrastructure being treated as essential infrastructure. This will be an important step to meet our wider ambitions around becoming a leading digital city region.

Other key recommendations related to standards and essential requirements for datasets themselves. This includes the need for datasets to be supported by clear documentation, clear dates for release, and comprehensive metadata. Datasets should also be released in non-proprietary formats that are acceptable for all. These findings are in line with the recommendations highlighted through our case study analysis.

Perhaps most importantly, this work has highlighted the key point that releasing data alone does not mean it will be used – we need to reduce barriers to use, as well as promote use. Our set of key deliverables for Greater Manchester, which includes continuous engagement with open data users, sets out some of the key steps needed to encourage greater open data usage.

Recommended deliverables for Greater Manchester

Essential deliverables for Greater Manchester

Our research has identified a series of deliverables for Greater Manchester to open more data for re-use. We have split these deliverables into three prioritised lists, based on feedback from data users and producers. The first list, presented below, consists of the essential elements – those that create the conditions for more data to be released. The wider list of deliverables, available in Annex A, aim to improve open data quality and quantity, and to create a good open data ecosystem in Greater Manchester. They are not essential for open data releases but are deliverables that Greater Manchester should or could implement.

There are 12 essential deliverables, grouped into four main themes: opening more data; setting standards and raising quality; inclusive governance and engagement; and embedding open data.

The recommendations are not prioritised – that is, there is no 'best' order in which to implement the recommendations. However, they have been designed to be iterative, with each new recommendation implemented both adding value to what has previously been implemented, and making further implementation easier. For example, theme 4 (embedding open data) is supported by inclusive engagement with open data both within and beyond the public sector, which is a facet of theme 3 (inclusive governance and engagement). Similarly, at an individual recommendation level, having consistent standards for datasets (deliverable 4) supports the release of more open data (deliverable 1), and vice versa. Therefore, while these 12 deliverables may seem ambitious, they can be approached individually or jointly, developed iteratively, and rolled out gradually.

It is also important to note that the list below should be constantly reviewed and updated, given that the demands of data users will change over time. The voices of users and producers of data should continue to shape the future of these deliverables.

Theme 1: Opening more data

The aim of this theme is to make open data more widely available and accessible.

Deliverable 1: Open more datasets, from more organisations, that users want and need.

The first step in any open data project is to make data open. Where necessary, consent and the provenance of the data needs to be considered. While the public sector does open data, some organisations open more than others. The Greater Manchester public sector should collectively open more data that is of value to businesses, academia and the third sector. This can help to grow demand and stimulate the view of Greater Manchester as an open, digital place.

Deliverable 2: Put in place a programme of work to identify, prioritise and support new data releases.

The skills, capabilities, and resources of public sector organisations in Greater Manchester differs. As such, it is important to put in place a programme of activity that identifies and prioritises new data releases, and then supports the release of this data.

Deliverable 3: Create an easily accessible place to share open data for each public sector organisation in Greater Manchester.

Every organisation should have a place to open its data. This could be a unique or existing location, such as data.gov.uk. If it is not a unique location, a list of all places where open data is hosted should be produced and made open.

Theme 2: Setting standards and raising quality

The aim of this theme is to improve the quality of information by setting standards. This includes the processes for ensuring these standards are agreed and delivered. It also aims to let people and businesses know what data is held, identify what data is available for re-use, and make accessing that data easier.

Deliverable 4: Develop and agree upon a set of basic data standards and documentation for the highest priority datasets.

As datasets are identified and prioritised, data standards should be developed with the data producer(s) and data user(s). This will ensure that there is consistency across Greater Manchester, adding value to the datasets and supporting their wider re-use. The data standard should also be published and supported by good, commonly agreed metadata and metadata standards.

Deliverable 5: Establish a central repository that lists and describes all datasets held by the public sector in Greater Manchester.

A list of all datasets would inform potential users of new datasets and help users to understand what data is available, and where from. This aim would be long-term and could be based on individual information asset registers.

Deliverable 6: Agree to open data in accessible, open formats.

Open data does not simply mean using an Open Government Licence. For data to be widely used, it needs to be provided in a format suitable for the users. Providing data to at least 3-star level, and in more than one format, can help support accessibility and ease of access.

Theme 3: Inclusive governance and engagement

The aim of this theme is to create inclusion in the open data process. This means including more organisations and individuals in open data governance and ensuring that the views of businesses and residents are included within open data decisions.

Deliverable 7: Establish an inclusive governance framework for open data across Greater Manchester.

There is currently no GM-wide governance of open data standards and approaches. To ensure that data releases are meaningful, it is important that all voices are heard. Existing governance structures could be expanded to consider open data, and to feed in the voices of business, academia and the third sector. Including both data producers and users in this governance framework can ensure data is as open as possible, and as closed as necessary.

Deliverable 8: Create an effective communication strategy that articulates the value and purpose of open data.

The value of data comes in its usage, not in its storage. Consistently emphasising this message is important to creating support for opening data. It is important to consider the messaging of open data – is it about simply being open, or is it about showing what Greater Manchester is, and can become? This message should be built on clear business cases and use cases.

Deliverable 9: Listen to and engage with data users, starting now.

The business engagement exercise was just a starting point of engagement between data users and data controllers. It is important that data users' voices are heard, and this is used to help prioritise data releases and actions. This deliverable would likely also include the need for a feedback platform for open data, creating a means to answer questions and accommodate feedback from the data reuse community.

Theme 4: Embedding open data

This theme aims to make open data 'business as usual', by making the case for more open data, and ensuring that public sector programmes and projects consider the value of the data they open at the outset.

Deliverable 10: Make the case for open data releases that will support business growth and innovation in Greater Manchester.

Our business engagement exercise highlighted the importance of data at a national level. Greater Manchester should encourage and support the release of more national datasets that allow local businesses, academia and third sector organisations to develop new goods, services, and research.

Deliverable 11: Ensure open data is embedded in all policy areas in Greater Manchester.

Open data affects all areas of policy work. However, there are stronger linkages with certain programmes of work, such as that of digital inclusion – for example, if people cannot access digital tools and services, they are unlikely to benefit from open data releases. The link between open data and other policy areas can be supported through the development of an open data policy, and the action plan outlined above.

Deliverable 12: Lay out an open data roadmap - an action plan - for the Greater Manchester public sector.

The roadmap should seek to deliver the recommendations in this report. It should also identify the individual actions that need to be taken by organisations to level up – from releasing open data initially, to increasing the quantity, quality, and usage of open data.

Further potential deliverables

A range of further deliverables is outlined in Annex A. These have been gathered from the case study review, workshop, and business engagement sessions. It is important to keep these deliverables under review.

The list of deliverables is split into three groups, using a MoSCoW type model: must have, or essential deliverables; should haves; and could haves. The second and third group build on the deliverables in the first group.

Examples of second group deliverables:

- **Design evaluation mechanisms to review the use of published data.** The outcomes of these evaluations should further support the essential deliverables above by informing relevant changes to the prioritisation and publication of data. Feedback gained from data users could also feed into this evaluation process.
- **Create a Greater Manchester-wide datastore.** This proposed deliverable would build upon the foundations laid above. A region-wide datastore would be valuable given that many of GM's residents and businesses work across borders. Our business engagement work suggested that GM might be an ideal scale at which to work on this.
- **Create a standards board for Greater Manchester.** One of the essential deliverables is to set out basic standards for the highest priority datasets. A GM standards board would build further upon these existing foundations, paying consideration to the trade-off between capacity and standards. Paying attention to data standards also further supports accessibility by ensuring data is easy to consume by the end user.

- **Engage with the open data community through hackathons or innovation days.** The outputs of these events could feed into the communication strategy outlined above. These events could also encourage or identify new and innovative uses of data.

Examples of third group deliverables:

- **Embed information governance team members within open data teams across the region.** Ensuring that each public sector organisation across the region includes an information governance team member would help to ensure that data is managed effectively. These IG leads would be able to support the application of the inclusive governance framework for open data.
- **Segregate data portal users to cater to multiple audiences.** Segregation would further ensure that data is provided in a format that is suitable to the users. This segregation could be applied to data portals maintained by individual organisations, or to a Greater Manchester datastore (as outlined in the second group deliverables above).

As the essential deliverables are put in place across Greater Manchester, it is important that the second and third priority lists are reviewed and used to outline further actions. By doing so, Greater Manchester can continuously improve, and respond to the needs of data users in the city region.

Related and future work

Alongside the project, the team has been involved in several parallel and related projects, including identifying future work. These relate to the work of the Local data Review but are not specifically part of the programme.

The project team commissioned two specific pieces of work around open data use and data quality in Greater Manchester. Through the course of the case study analysis and business engagement, it became clear that these pieces of work would support future open data activities, rather than essential open data deliverables found in the previous section. The remaining areas of related work outlined below have emerged directly from the case study work and business engagement sessions.

The projects reported here are further areas of work and actions that will contribute to the creation of a good open data ecosystem in Greater Manchester.

Data Quality Analysis

In early 2020, GMCA asked Norman Paton from the University of Manchester to explore a range of issues relating to data quality, specifically focusing on the ease with which public sector open data can be reused. Common data quality issues were identified and illustrated with practical examples.

This review highlighted the inherent tensions between data quantity and data quality. An “open by default” approach can be intuitively appealing, but this needs to be balanced with the need to ensure adequate quality, consistency and usability across multiple datasets and sources. At the same time, a focus on data quality should not come at the expense of getting more data open. The balance between quantity and quality should be carefully considered and evaluated, with iterative amendments to the approach over time.

The findings of this data quality analysis will support our wider work in understanding and improving data quality for public sector open data.

Future potential actions:

- Create a data quality checklist for public sector organisations to use.
- Provide guidance to organisations to support them in using the data quality checklist.
- Review a sample of released data against the data quality checklist to monitor its effectiveness, and continue to offer support to data producers to improve open data quality.

Public Consultation

The focus of our Local Data Review is around potential reuse of public sector data by other organisations, primarily the private sector, to support innovation and productivity growth. Nevertheless, the public's use of open data is also a key consideration for how Greater Manchester approaches the release of its information.

Parallel to the Local Data Review, the team has conducted a public consultation to find out whether residents of Greater Manchester use open data, how they use it, and their thoughts around open data. We will analyse the public survey findings in more detail and consider the impacts of this consultation on what data we release, and how it is released.

Future potential actions:

- Gather the survey responses and analyse the findings to identify key themes and important messages from the survey.
- Build the survey analysis into the wider action plan for Greater Manchester, to ensure that the open data needs and preferences of the public are included.
- Continue to engage with members of the public through a range of communication channels and approaches.

Transparency

Linked to the above, it is important that the role of transparency remains central in the consideration of open data releases. Our business engagement work suggested that electoral, transparency and performance information was still valued. As an open data action plan is developed, the work should look to encompass the transparency agenda, aiming to release information that is valuable from an open government perspective.

Future potential actions:

- Collate a list of electoral, transparency and performance information held by the public sector in Greater Manchester.
- Outline whether each set of information in this list is already open or could be made open, linked to a range of defined standards.
- Consider how this information might be included in a prioritised list of data released in Greater Manchester.

Health data accessibility

The business engagement work conducted by Open Data Manchester highlighted a series of asks from academics and businesses around the use of health data. In addition to requests for releasing data, there were also suggestions for enhancing the use of health data. These included the development of health data ‘sandboxes’ for users to utilise the data in a safe and secure environment; and the creation of NHS Data Managers to promote the use of anonymous health data. Some of these suggestions fall outside of the remit of this work, as they are not focused on making health data open. As such, it is important these ideas are followed up with the appropriate organisations and teams.

Future potential actions:

- Provide detail on the findings to relevant stakeholders, in particular the suggestions that GM:

- Set out comprehensive information governance processes around health data, which include clear procedures on how to obtain approvals for access to health data.
- Consider building a secure platform or health data ‘sandbox’ for users with approved access to analyse data that cannot be freely shared on the internet.
- Explore the viability of creating synthetic data for sensitive and protected data.
- Develop best-practice guidelines regarding data anonymisation and other privacy-enhancing techniques.

Open data process in GMCA

A parallel process to the Local Data Review was the aim of the GMCA itself opening data. Several potential datasets have been identified, and it is proposed that the team explores the process for releasing data within the GMCA.

Future potential actions:

- Develop appropriate processes and governance for releasing more open data in the GMCA.
- Open more basic location data – such as the location of all waste sites and the types of waste they recycle.
- Open more asset data – the location of fire and rescue service stations, alongside GMCA land and property assets.
- Explore the possibilities of opening alternative datasets – such as data on the location and recordings of air quality sensors around Greater Manchester.

Next steps

The section above sets out the deliverables to open more data in Greater Manchester. These need to be delivered. It is recommended that partners across Greater Manchester work on the following actions:

1. Develop an action plan

One of the essential deliverables above sets out the need for an action plan. It s important that there is continued involvement of key partners across the region, such as members of the Local Industrial Strategy Programme Delivery Executive, members of the Digital Portfolio Delivery Executive, and members of the Information Board, to agree the next steps for the region. This step will require obtaining a mandate for the delivery of the recommendations and will include assigning responsibility for different actions to relevant teams and organisations.

2. Summarise and incorporate the public consultation

The aim of this Local data Review has been primarily about the use of public sector data for productivity growth and innovation. However, an important driver of the open data agenda has been transparency, and public access to information. The public consultation on open data in December 2020, conducted alongside this review, focused on the public's use of open data. It is important that the findings from this consultation are incorporated into the city region's wider plans for opening data.

3. Develop further work on a data quality checklist to support data releases

It is critical to ensure open data is of high quality. Poor quality data can erode trust, as well as create erroneous outcomes. This work should seek to provide a checklist for data producers to consider before releasing data.

4. Ensure open data is reflected in relevant strategies and programmes

A number of strategies and programmes are in development, which have a key link to open data and its use. It is important that an open data approach is particularly reflected in both the Information Strategy and the refreshed Greater Manchester Strategy.

5. Open more data

The work has identified a range of simple datasets – several outlined above – that could be opened up for re-use. GMCA and partners should explore these opportunities to understand which of these datasets can be prioritised and opened on a consistent basis.

Appendix 1 – Note on terms and definitions

Throughout this document, a number of terms are used that it would be helpful to identify here. The below provides a few details on key terminology for the report:

Data users: data users refers to all organisations or individuals that access and use information. In this report, the term applies primarily to non-public sector data users. The focus of this report is on businesses, particularly digital and data businesses. As a result, 'data users' in this report references non-public sector data users, primarily businesses.

Data owners: this term references an array of organisations and individuals involved in the production, ownership, publication, or control of data. There are a number of legally defined roles that are included within the use of this term. Generally, this term aims to cover those involved in the supply side of information, rather than the demand / use side.

Open data: in this report, the term open data refers to data that aligns with open definition: "Open data and content can be freely used, modified, and shared by anyone for any purpose". For the public sector, this would largely involve releasing data under an Open Government Licence.

Appendix 2 – Open data project development timeline

The starting point is simply **opening data**. The way this approach is set up can affect decisions in future.

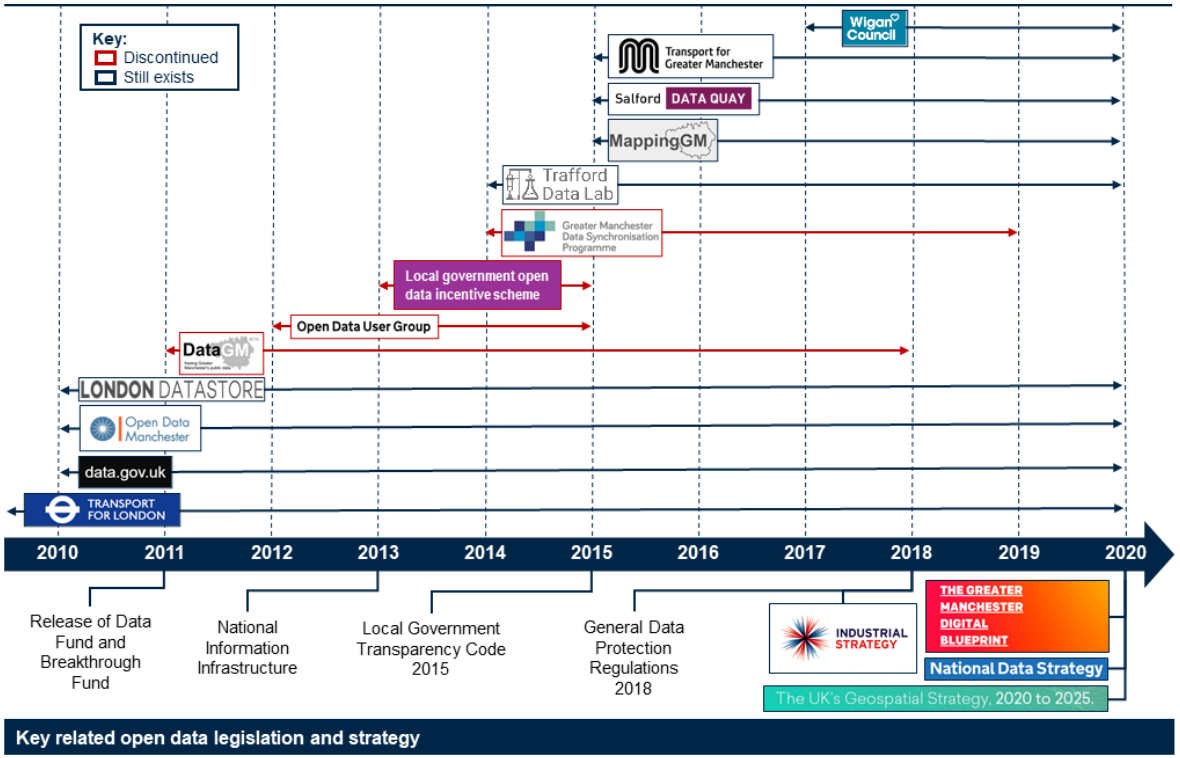
The focus often shifts to increasing the amount of open data. As more data is opened, more robust **data infrastructure and architecture** tends to be needed.

As more data is opened, there is the possibility of both **data divergence** and **portal divergence**. A focus on improving data quality becomes important.

The future of open data looks set to move towards greater **data visualisation** and greater **open analysis**, rather than simply open data. This requires an understanding of how data is used.

Source: GMCA Research, 2020

Appendix 3 – Timeline of selected local and national open data initiatives

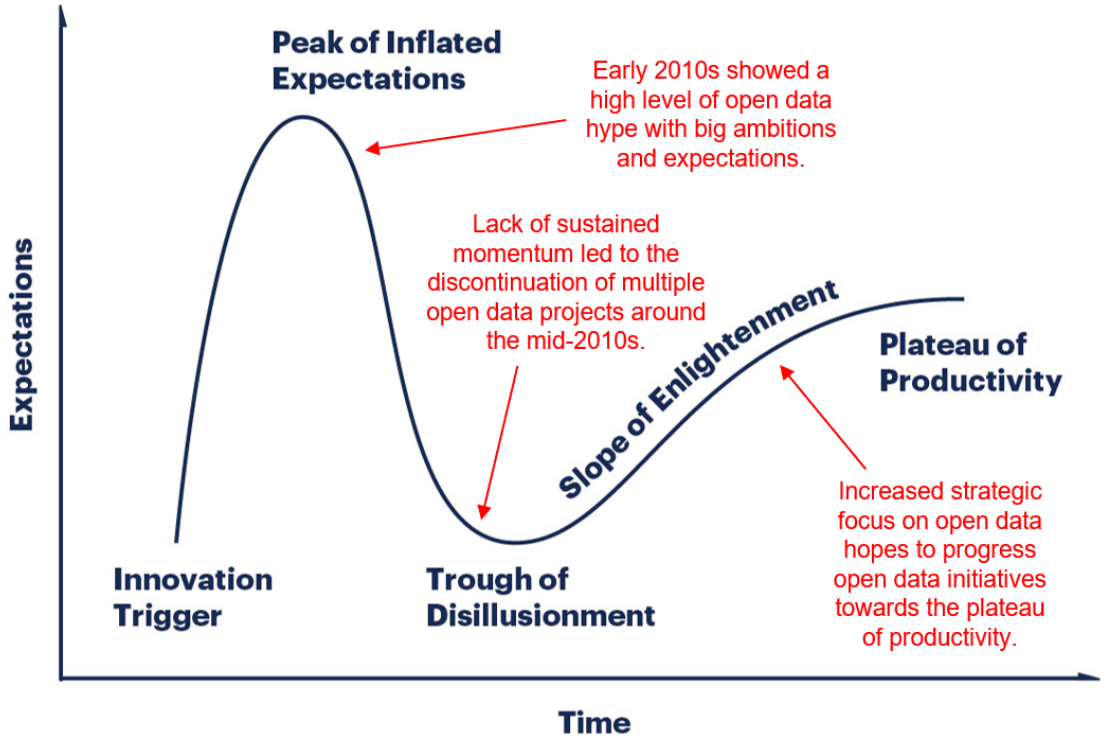


The image above shows a timeline of selected local and national open data initiatives from 2010 to 2020, alongside some key legislation from the same period. Some innovative projects, such as DataGM and the Greater Manchester Data Synchronisation Programme, were arguably ahead of their time. In this respect, they may have been too ambitious, and have ended up being

discontinued. Others that have taken a more gradual approach to opening data, such as data.gov.uk and Trafford Data Lab's open resources, continue to exist today.

The timeline broadly shows that high initial enthusiasm for new open data projects towards the start of the decade was followed by the subsequent discontinuation of several projects. More recently, there has been a renewed wave of open data activity, and open data has risen on local and national strategic agendas. This pattern broadly ties up to the common progression of new technology outlined within the Gartner Hype Cycle. Appendix 4 expands upon the Gartner Hype Cycle in more detail.

Appendix 4 – Gartner Hype Cycle, with a focus on open data



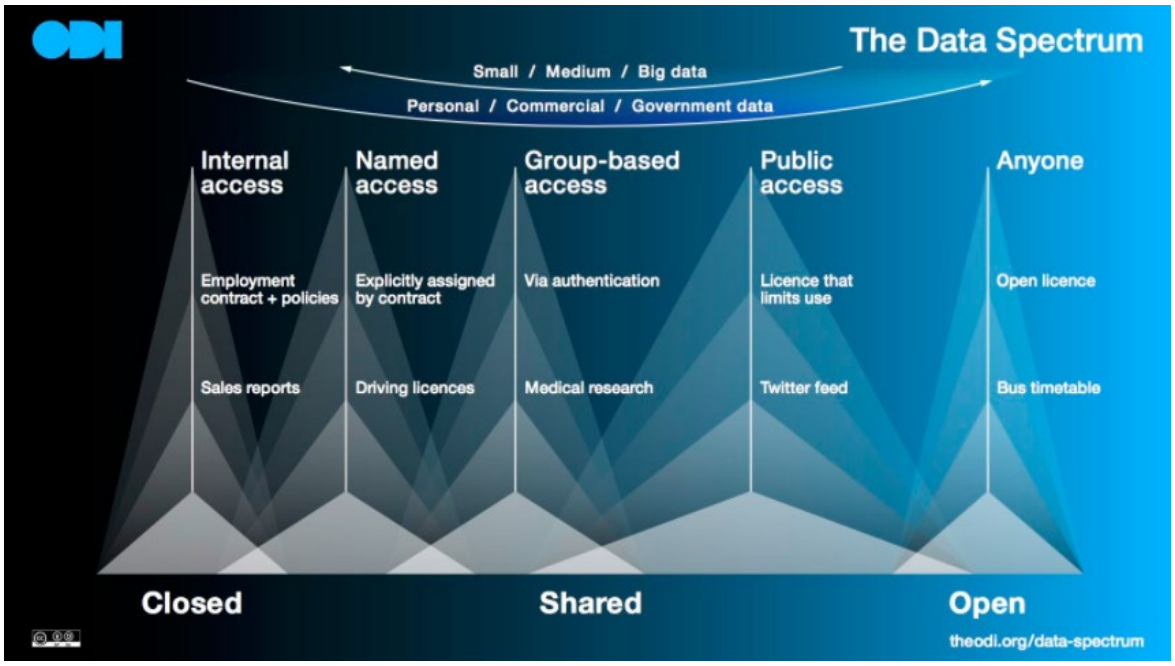
The Gartner Hype Cycle provides a graphic overview of common patterns of maturity and progression in the adoption of new technologies over time. Trends in open data over the past decade have broadly correlated to the stages of this cycle, both within Greater Manchester and the UK.

Enthusiasm for open data gained significant momentum around 2010, leading to a 'peak of inflated expectations' and a flurry of new open data projects and initiatives at this time. However, in the subsequent few years some of this momentum was lost and some of these programmes were terminated, as highlighted in Appendix 2 above. This dip is a common pattern in the maturity of new technology based on the Gartner Hype Cycle. However, a range of additional factors contributed to this loss of momentum. For example, key open data projects of the time, such as data.gov.uk, did not meet initial expectations for user numbers or datasets.

Despite the loss in momentum in the mid-2010s, there has been renewed strategic interest in open data more recently. This has coincided with rising awareness of the

commercial opportunities and enhanced productivity generated through open data initiatives. Renewed interest has arguably pushed the open data agenda up the ‘slope of productivity’ towards a ‘plateau of productivity’. At this more mature stage of progression, open data now remains high on local and national strategic agendas and is widely understood to be an asset for the UK economy.

Appendix 5 – The Data Spectrum



Source: The ODI

The ODI data spectrum shows the full range of data from closed to open. This highlights the wealth of data that could be made open. It also shows what data can be considered closed or needs processing to be made available as open data.