Discovering the future of the London Datastore

Findings and recommendations from a discovery project. December 2019

Open Data Institute

Executive summary

As the London Datastore prepares to embark on its second decade, is its organisational and technical set-up still fit for purpose? And is it still the best vehicle to equip agencies, local authorities, companies, organisations and Londoners to get the best out of data, to inform better policies and decision making, to enable the creation of better services and innovation, and tackle the big challenges of the years to come?

Those are some of the questions the Open Data Institute set out to answer by engaging with hundreds of people and organisations on a three month discovery project, on behalf of and in close cooperation with the Datastore team at the Greater London Authority.

Through research, interviews, workshops and a survey, many insights emerged about the needs of data stewards and users, and the potential of enabling better access to high-quality, relevant and timely data. Some of those insights, such as a need to improve the findability of datasets so that people can find the data they need, are similar to the issues other stewards of data portals and data platforms are facing and tackling at the moment. Others, like the need for increased coordination across the many agencies, local authorities and other data stewards in and about London, are more specific to a team and platform aiming to create impact for the millions living in, working in or visiting the UK's capital city.

Based on those insights and existing research, we recommend six actions for the Datastore team — some quite tactical, others more long term and visionary — across three themes. The themes are
- making the London Datastore a better source for data,
- creating a destination for insights, and
- being a trusted guide and steward to the data community.

We recommend to:
1. Improve the findability of the data
2. Increase the variety and volume of data on the Datastore
3. Showcase data reuse
4. Document best practices
5. Champion standards adoption and development
6. Encourage and facilitate collaboration

About this report

Introduction

Context

The London Datastore is a data-portal pioneer launched in 2010. It is a platform where anyone can access public data relating to London. Its second iteration in 2015 – which greatly improved its design and functionality – won an ODI award[1] in the open data publishing category.

As a model for access to data, portals have proven their usefulness. Whether they are national, local or regional, thematic or sector-focused, they have been empowering people, increasing transparency, and enabling innovation. Many of them, however, were set up without sustainability in mind, and are now sitting unloved and underused.

The open data portal model has also shown its limits. It is now well understood that data sits on a spectrum between closed and open, and that data stewards can unlock value by increasing access to data in ways that will maximise its value while minimising potential harmful impact. The London Datastore team has taken steps in this direction already by enabling secure sharing of data on the platform.

Portals also need to evolve with technology. The past 10 years have seen a rise in application programming interfaces (API)s, 'knowledge graphs', improved dataset search, and an increasing use of live and streaming data. The ability to enable the discovery of, and access to, a range of data sources will only become more important in the years to come. Adapting to this technological change will require independent, trustworthy data governance.

Motivation

Defining the shape and scope of the future London Datastore is challenging. The Datastore must respond to the needs of its current users while becoming fit for purpose for the future.

A discovery phase – exploring who the users are and what they need, what data, governance and technology exists, and what the requirements of the product may be – is a valuable investment. It can radically speed up future development by providing clarity and reducing the potential waste of developing unnecessary features, or even products which don't meet user and market needs.

Our approach for this discovery phase combined mixed-method research activities and a collaborative, iterative approach, to meet the following objectives from the Datastore team:

1. *To understand the wider service that the London Datastore supports and what additional positive contribution it could make.*
2. *To understand the needs of Datastore users and potential users and the intersection between this and your ambitions.*

---

[1] Andrew Collinge (2015), 'The Morning after the Night before: international recognition for the London Datastore',
https://data.london.gov.uk/blog/the-morning-after-the-night-before-international-recognition-for-the-london-datastore-2/

3. *To define the mission and vision for the London Datastore.*
4. *To understand the steps needed to implement the next iteration of the London Datastore.*

Methodology

Our approach combined the expertise of the ODI in data technology and policy, as well as creating new data access approaches such as data trusts, with a robust, mixed-method research plan. We conducted the following research activities:

Activity	Purpose
Desk research	Learn from the wider data ecosystem, building on existing work from ODI and others.
Four expert interviews	Tap into expertise from the wider ecosystem; part of a collaborative approach. Three interviews were focused on data stewards sharing/publishing, with one interview focusing on reusers.
Publishing-focused practitioner workshop	Understand the experience of data stewards publishing or sharing data to the Datastore or similar platforms. Identify issues, barriers to effective publishing and infer tactical remedies for the Datastore; understand needs of data stewards to inform longer-term direction.
Data reuse practitioner workshop	Understand the experience of people and organisations accessing and using data published or shared on the Datastore or similar platforms. Identify issues, barriers to data discovery and use, and infer tactical remedies for the Datastore; understand needs of data users to inform longer-term direction.
Workshop with members of data teams in London boroughs (Organised with London Office of Technology and Innovation)	Similar in scope to the publishing-focused practitioner workshop, but with a focus on the needs of data teams in London boroughs.
Public survey	Gather broad qualitative input from users and non-users of the Datastore. The questionnaire was mainly focused on the needs and experience of data users.

We ran the project in close cooperation with the GLA team to ensure that the GLA team could get firsthand insight by being present in interviews and workshops. The direction of recommendations was also tested and refined through a vision workshop with GLA team members.

Findings and recommendations

This section brings together findings from all the activities in this discovery phase, focusing on actionable insights as much as possible. A more comprehensive summary of the desk research, workshops, interviews and survey can be found in appendices [1], [2], [3] and [4] respectively.

London was one of the first cities[2] to recognise the importance of access to data and insights as a means to improve decision-making, increase transparency, and enable innovation. In a world of shifting political and economic priorities and of fast technological change, the key to the continued success of this approach is to enable people to find, access and use data about London in a way that is not overly dependent on current administrative boundaries or particular platforms.

Defining the future role of the Datastore involves recognising that providing a data portal is just one aspect of building a stronger data infrastructure for the city. A portal is a method of publishing and sharing data for those that need it, while offering a means to discover a wider range of datasets. Making the portal a key contributor to building a stronger data infrastructure means building a community around it and developing guidance, standards and best practices for users. Improving the interface of the platform itself will help to facilitate better discovery and use of data, but creating an open, trustworthy data ecosystem for London will require a broader set of activities.

We review the Datastore as **an exemplar in transition**, summarising the many things it does well and should continue doing, before presenting our recommendations in the following sections. Those recommendations focus on becoming a **better source of data**, creating a **destination for insights** and becoming a **trusted guide and a steward.**

These three elements are not mutually exclusive: the GLA should pursue all three in order to meet the varied needs of the publishers and users it is supporting. Based purely on the findings from this Discovery project, our recommendation is to prioritise becoming a **better source of data** and a **trusted guide and a steward.** Investing in a **destination for insights** would be a useful strategic decision.

GLA should clearly define and communicate its direction, in order to manage the needs and expectations of the ecosystem of data stewards, users, and other people affected by the collection and use of data in London. Just as Berlin[3], New York[4] and a group of six Finnish

---

[2] The Datastore was created not long after the concept of Open Data was codified, according to Emer Coleman (2013) 'Lessons from the London Datastore' in 'Beyond Transparency', https://beyondtransparency.org/chapters/part-1/lessons-from-the-london-datastore/
London was one of the first in Europe, according to European Data Portal (2016) 'Open Data in Cities' https://www.europeandataportal.eu/sites/default/files/edp_analytical_report_n4_-_open_data_in_cities_v1.0_final.pdf
[3] CITTEGO (2018) 'Berlin Open Data strategy', http://www.citego.org/bdf_fiche-document-1195_fr.html
[4] NYC Open Data (2019) 'Open Data for All Report', https://opendata.cityofnewyork.us/wp-content/uploads/2019/09/2019_OpenDataForAllReport.pdf

Cities[5] have published an open data strategy, London should clearly set out its vision about increasing access to data across the data spectrum.

An exemplar in transition

As one of the first city-level open data portals, the London Datastore has for the last ten years, been an exemplar for other cities.

It is evident that over the years, the Datastore has become an important part of London's data ecosystem, with around 60,000 people a month using the portal and over 4000 datasets available for download under an open licence — with another 2000 datasets shared on the platform. The Datastore team has updated the portal to meet changing user needs by trying to improve the experience for technical users, while also recognising that users who have a less technical background may find visualisations and analyses more useful than the raw data.

The research workshops suggested that the platform itself does a good job as an open data portal: publishers know how to upload to the portal and users can generally find data when it is available. The core features all were useful to part or all of the people we engaged with: access to open data; the creation and curation of insights and visualisations; and access to shared data too.

With that said there is room for improvement, particularly around search and navigation as well some of the functionality around sharing non-open data. The ability to iterate will be essential for ensuring the Datastore remains fit for purpose as its scope evolves from an open data repository to a central registry of London's data.

Many of the difficulties that people face in using the Datastore cover a range of cultural and process barriers. The majority of the current technical barriers, some of which are outlined in the following sections, would not be too complicated to overcome.

Our research shows that currently people have low expectations of the Datastore as a technical platform. It needs to ensure that data is findable, usable and linked to related data and documentation: the main purpose of the Datastore is to be a trusted catalogue, rather than store or platform for data.

The GLA have begun taking steps to ensure that the Datastore is able to facilitate access to data across the data spectrum — providing access to open data and encouraging publishers to be as open as possible, and also offering a platform for more controlled sharing of data. We believe that this role, which will help to overcome potential barriers and mitigate risks around sharing of data will only become more important, ensuring the relevance and sustainability of the Datastore in the future.

A better source of data

Our research suggests areas where current features of the Datastore could be improved. Many of them can be addressed in a way that makes the Datastore a better source of data:

---

[5] 6aika 'How does it work?' https://6aika.fi/en/what-is-6aika/how-does-it-work/

improving the findability of data, and **increasing the variety and volume of data covered by the Datastore**.

Recommendation 1: Improve the findability of the data

One of the core requirements of a data portal is to help people find the data they need.[6] The existing search function of the Datastore allows you to search by topic, publisher name, format and geography, but users still find they have to 'click around' a lot and, when they do find the data they need, it can be hard to find it again on returning to the portal. Finding datasets can be quite difficult if you don't already know exactly what you're looking for.

Understanding complaints about search and navigation

Our research repeatedly showed a need for improved quality and consistency of data and metadata available on the Datastore. This can, in part, be addressed through a focus on increasing the findability of the data across all modes of discovery, including browsing and searching.

"[What would make the Datastore better] is an easier retrieval of data and search function"

"Easier navigation to specific data sources"

Participants of the public survey

Users identified issues with navigation and search as one of their major complaints with the Datastore. Navigation and tagging are common problem areas for data portals, and it can be hard to keep up with technological advances in search functionality. The City of New York has mitigated this problem by providing simple guidance about how to use and navigate the platform for both new and existing users.[7] The City of Amsterdam has gone one step further and embedded clearly-defined search features, and provided a mechanism for users to feed back about it.[8]

Addressing concerns about "outdated" data

There is a perception that data on the Datastore is 'outdated', possibly in part due to the limited information available about update frequencies. Unclear update frequencies can negatively affect whether the data will be used to create impact. As one user noted:

"We really need to know how often the data will be updated
before we can really commit to use the data".

This issue is exacerbated for some users by a lack of feedback and engagement mechanisms, which they feel may have been useful as a way to ask whether a dataset is up to date. The analysis section of the platform contains links to descriptions of various GLA teams, but no

---

[6] European Data Portal (2017), 'Recommendations for open data portals: from set up to sustainability', https://www.europeandataportal.eu/sites/default/files/edp_s3wp4_sustainability_recommendations.pdf
[7] City of New York (2017), Open Data Portal, 'How to - Getting Started With Open Data', ://opendata.cityofnewyork.us/how-to/
[8] Amsterdam Data and Information, https://data.amsterdam.nl/artikelen/artikel/contact/c8f4d1da-75f3-4ee7-93e8-256b201d6ccf/

clear information on how to engage with them. This can make it hard for users to contact relevant teams with questions, especially when most of the data are from stewards outside of the GLA.

We recommend a review of how the information on timeliness of datasets is presented throughout the Datastore, with an emphasis on whether a dataset is the latest available: making it clearer when a dataset is the latest version available, regardless of the date at which it was issued. Conversely, datasets superseded by others should be clearly marked as such, and corrections issued (see for example how the Office for National Statistics regularly releases and updates its key datasets) should be given prominence to increase trust in the relevance and timeliness of data presented on the Datastore.

Additional recommendations on engagement, guidance and access to data through APIs (which are perceived as being always up to date) are addressed in subsequent sections.

Reviewing categories for better navigation
For a better browsing experience, we recommend reviewing the categories used for navigation on the Datastore. Participants of the public survey pointed out categories that were relevant for them and weren't listed on the Datastore – such as education, culture or population statistics.

Exploring and testing alternatives with users and publishers, and being open to the categories and tags used for navigation being a fluid and evolving set, should be a first step towards better navigation.

The Datastore already provides useful onward journeys with links to other datasets by the same publishers and other related datasets, which ought to help with navigation, especially for users landing on a dataset page from search results. There may be value in testing the usability of those navigation mechanism, and the perception of the relevance of the links to related datasets.

The main focus towards better findability of datasets on the London Datastore should, however, be about metadata - information about the datasets including topic tagging and descriptions.

Providing curated guides
Content curation is, alongside browsing and searching, one of the major modes of discovery, but one which is often underused in the digital space, where the emphasis is on indexing and categorising large catalogues to help people find what they know they are looking for rather than providing guidance and curation to help them find what they need.

Just as we need well organised data catalogues and improved search facilities, we need librarians for data catalogues[9]. Curated guides oriented around common use cases can be a cost effective mechanism for increasing findability of data (alongside, and possibly ahead of, efforts to increase metadata across hundreds of datasets).

_____

[9] Joyce L. Ogburn (2010) 'The Imperative for Data Curation', https://muse.jhu.edu/article/378049

A good curated guide can point to examples of reuse if they are available (see Recommendation 3: Showcase data reuse), but a guide oriented towards "here is the data you need for this challenge and how to use it" can exist before there are any examples of that use.

Reviewing metadata for a better experience and more effective search
Poor and inconsistent metadata is one of the major issues highlighted throughout our findings. One insight from user engagement is that it can be hard to understand the scope of a dataset and any limitations that might inform its use. Often there isn't enough information about each dataset to help users to find out if it might be useful to them or not.

> *"What [our] organisation needs is to be able to understand that data in some depth [such as] geographical depth, a sort of day to day time specific depth, etc […] And also understand how reliable in itself that dataset is. And that's probably why [specific] business would typically look to other sources"*
>
> – User of open city data

Better metadata should increase the quality of search results[10], both for the internal search engine of the Datastore and public search engines, which are increasing in importance for users. In a blog post published in April 2019[11], the Ordnance Survey confirmed this trend, writing: *Our user research also revealed that 75% of users start their search for geospatial data by using Google.*

General web searches for data[12] will improve over time as search engines increasingly analyse the content and structure of datasets. In the short term however, they will continue to rely heavily on good quality declarative metadata.

A first step to address this issue should be to audit the quality, completeness and consistency of metadata created for datasets on the Datastore, review all documentation and guidance provided for publishers, and decide what minimum threshold of metadata richness and quality would be a minimum to support the needs of users.

The London Datastore currently fares badly in the amount and granularity of machine-readable declarative metadata for the open datasets it links to. Using tools such as the Google Structured Data Testing Tool[13] to improve results for London Datastore datasets, and benchmarking against other sites such as the Office for National Statistics' or the French Government's open data portal should yield fast and measurable improvements in findability of the Datastore's datasets.

---

[10] Koesten, Simperl et.al. (2018), 'Everything you always wanted to know about a dataset: studies in data summarisation', https://export.arxiv.org/abs/1810.12423
[11] Geospatial Commission Data Discoverability – making geospatial data easier to find (2019), https://www.ordnancesurvey.co.uk/blog/2019/04/geospatial-commission-data-discoverability-making-geospatial-data-easier-to-find/
[12] Chapman, Simperl, Koesten et.al (2019), 'Dataset search: a survey' https://export.arxiv.org/abs/1901.00735
[13] Google's Structured Data Testing Tool, https://search.google.com/structured-data/testing-tool

This recommendation for better metadata applies equally to both open and shared data catalogued in the Datastore. While there are occasional privacy or sensitivity concerns, documenting the existence of data, even if access is restricted, is generally safe. Cataloguing data across the spectrum can help drive discovery and increase reuse, provided there is also a clear and transparent way to request access.

This focus on increasing the findability of data, regardless of whether it is stored and available on the datastore itself, would lead to the Datastore becoming more of a **central registry** facilitating access to data than a store of data.

Recommendation 2: Increase the variety and volume of data on the Datastore

In addition to its original remit around providing access to public and open data, the Datastore has increasingly been used to provide and manage access to shared data.

Engaging with the community to prioritise which data to add

Understanding user priorities and using this to identify data to publish can be hard, and the majority of data is shared by borough Councils in response to a legal or statutory need, rather than being user driven.[14]

> *"We want to make sure that the time and effort spent to publish the data*
> *Is relevant and useful for others."*
>
> –Participant on the public survey

The publishers we talked to would like to better understand the needs of users of the London Datastore. Insight into who users are and what they do with different types of data could help publishers to prioritise the data they provide, and the way they provide it.

The Datastore team should explore mechanisms to increase user engagement and input into what data (and insights) ought to be made available.

The most effective approach is a problem- or challenge-based one: working with data users and people affected to identify a problem which data can help solve, and then increasing access to the data required to solve it. Focusing on challenges rather than simply creating inventories of data is more likely to yield reuse[15].

The London Datastore already does this in some cases: work on the Night Time Observatory[16] will bring together boroughs (users) and GLA (publishers) to publish data on the night time economy to aid the boroughs in creating their night time strategies.

---

[14] GovEx Labs (2019), 'Open Data Portal Requirements', https://labs.centerforgov.org/open-data/portal-requirements/
[15] Open Data Charter (2018), 'Publishing with Purpose', https://medium.com/@opendatacharter/publishing-with-purpose-introducing-our-2018-strategy-ddbf7ab46098
[16] London Night Time Commission (Jan 2019) 'Think Night: London's Neighbourhoods from 6pm to 6am', https://www.london.gov.uk/sites/default/files/ntc_report_online.pdf

A recent report by the European Data Portal[17] showcases a number of ways national data portals across Europe have been involving a broad community in prioritising what data to collect, publish and make available. Once set up and well taken care of, such a community can be an invaluable asset in understanding where issues lie — in particular around data quality — and get a much better view of the return on investment for the platform.

See also Recommendation 3: Showcase data reuse for another way to address the need for publishers to better understand what data users need and what they do with the data.

Enabling more technical means of access
Another dimension which should be considered in expanding the variety and volume of data is about the variety of technical modes of access.

For example, many data users who participated in this research were used to accessing data through query APIs or streaming APIs, especially for real-time sensor data, and were critical of a platform built mainly around static datasets.

"Because of the business we're in, technology is changing. [...] we are switching from static to real time and predictive data"

–Publisher of city data

"Most of [London Datastore data] is fairly static. Transport API requires considerable technical expertise to engage with it"

–Participant on the public Survey

We recommend considering this feedback in planning the future of the Datastore, but with nuance. In early 2018, an ODI team wrote about the experience of working with open data from the perspective of a private sector application developer[18] and concluded that dual access to data — via both APIs and the availability for download of frequent snapshots of the data, was often preferable to only providing APIs or only static data.

Continuing the exploration of alternative access models
The London Datastore should continue in its efforts to increase access to data across the whole spectrum from shared to open. There is a significant role for the GLA to play in exploring new models for access to data — such as data trusts, which the GLA was a pioneer in piloting —, new institutional approaches to stewarding and managing access to data, and working to support sharing across a wider group of organisations.[19]

It is important to recognise that there are a variety of models for increasing access to data (see for example, The ODI's Data Access Map[20]). Rather than adopting a one-size-fits-all approach the GLA and its partners should test which models work best to address specific challenges,

[17] Open Data Portal (2018) 'Open Data Maturity in Europe',
https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n4_2018.pdf
[18] Open Data Institute (2018) 'Prototyping with open sports data'
https://theodi.org/article/prototyping-with-open-sports-data-report/
[19] Open Data Institute, 'Mapping the wide world of data sharing',
https://theodi.org/project/the-data-access-map/
[20] Open Data Institute (2019) 'The Data Access Map', https://theodi.org/project/the-data-access-map

and ensure these models can be catalogued and/or accessed via the Datastore. Exploring new models should include a focus on building strong governance around data, rather than merely investing in technical platforms.

For the GLA's own work, data access needs vary from project to project and new solutions may only be required temporarily. When exploring the infrastructure needed to enable new access models, the Datastore should use open source alternatives where possible, rather than creating new software or relying on long term agreements with suppliers.

A destination for insights

The Datastore already offers narratives and insights, which is one of the ways it caters for a diverse audience of technical and non-technical data users, creating value from the data held in the Datastore, and making a case for the use of that data.

Highlighting insights, but not at the expense of data access
Insights from the research confirmed the need to walk the tightrope of aiming to be useful to a very diverse set of users, while not failing to serve them by trying to be 'everything to all people'. Analysis and insight are time consuming, costly, and rely on different skills to those required to manage an effective source of data.

While the two objectives of creating valuable insights and giving effective access to data are not mutually exclusive, the focus on one could affect budget dedicated to the other. Too much importance given to generating insights from data may distract attention from the necessary work of improving how it can be accessed, used and shared.

Although GLA needs a space to host insights to address its own needs, responses from the public survey show that 75% of participants typically access information about London by downloading data and only 50% does it through by accessing ready-made insights. We would recommend engaging further with GLA data science teams developing data insights to understand their specific needs when publishing insights (and data) from specific projects.

Recommendation 3: Showcase data reuse

There is undoubted value in showcasing usage of the data, both as an exemplar for future users and as a tangible demonstration of some of the impact created by increasing access to this data.

Working with a community of data users
This does not have to be done through a heavy investment to generate all the insights and analysis of the data in-house, but instead can benefit from a collaborative, community-focused approach.

The blog section currently on the Datastore platform provides an opportunity for people to share their stories, experiences and calls to action concerning data about London, however the range of content listed is currently very broad and difficult to navigate which may limit the impact and insights that can be drawn from it.

" [The Datastore would be more useful] with show and tell examples of how people use it."

By working effectively with the community of data users, the Datastore would not only host analysis and insights generated by its internal team, but should also showcase interesting uses of the data. As for the generation of insights and analysis, this curatorial work can be done by a dedicated team, or it can leverage the enthusiasm of the community by encouraging data users to showcase their work on the platform. The open data platform of the French government, data.gouv.fr, includes a reuse section where community members can document their use of the data on the platform, and thus provide examples of the many ways each data set can be useful. Several examples of successful innovative data reuse in city portals are listed in the Desk Research Summary.

Such an approach is likely to become the norm: the 2018 Open Data Maturity in Europe[21] report from the European Data Portal highlights that "21 portals (81%) have a designated area to promote Open Data use cases" (and in 20 of these cases, the portal allows reusers to upload their reuse examples) but also notes that for now, "National Open Data portals seem to be reluctant to enable the broader involvement of the Open Data community on the national portal".

This kind of community-generated information does not, obviously, mean free content: the platform team moderates submissions, and provides high quality documentation, guidance and tools for the community to effectively curate and document examples of data use. Building expertise in collaborative maintenance of information through a community can be a challenge but guidance is available from the ODI[22] and others[23].

A trusted guide and steward

This third theme was the one most consistently addressed across all discovery activities in this process. There was a near-unanimous demand for quality and consistency, and, especially, for guidance and standards from publishers.

> *"I fully trust the Datastore on security, privacy, etc.*
> *I am not sure about the quality of third-party data, how consistent it is, etc"*
> –Participant on the public survey

That focus is not typical of what data portals or platforms offer. Providing extra help and support to data stewards and users, convening communities and leading on standards is not, strictly speaking, necessary for the efficient operation of a data portal or platform, but it can create the right conditions for the platform to reach its potential. In other words, standards, guidance and best practices are part of building a stronger data infrastructure, but they are often overlooked due to a focus on technical platforms.

---

[21] Open Data Portal (2018) 'Open Data Maturity in Europe', https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n4_2018.pdf
[22] Open Data Institute (2019) 'Collaborative Data Patterns' https://collaborative-data.theodi.org/
[23] See e.g. Wikidata https://www.wikidata.org/wiki/Wikidata:Community_portal

To increase its role and standing as a trusted guide and steward, the London Datastore and the GLA should address the following key areas, which make up our final three recommendations.

Recommendation 4: Document best practices

Documentation and highlighting best practices gives guidance to data stewards and users, it also shows them what 'good' looks like.

Clarifying what 'good enough' looks like

Not knowing acceptable thresholds of quality and what constitutes excellent to aim for is, as documented in the 2018 ODI report 'What data publishers need'[24], one of the main causes of anxiety and paralysis for open data publishers.

An earlier recommendation focused on increasing the quality and consistency of metadata. Better guidance for data stewards on the platform will not only help with this goal, it will also address a concern we observed in our engagement with publishers: there is a perceived absence of guidance to tell individuals how to describe data (metadata), how to name fields within the data or what 'quality data' looks like. There is no clear minimum standard for publishing to the Datastore. Guidance and training can help individuals be confident that they are doing the right thing.

Providing guidance for data users

There is a recognition that the quality of available data varies, and the absence of contact points makes it hard to raise concerns about data quality with publishers.

Users want better guidance and communications as a priority. Examples from Vancouver, who have produced a guide for new and advanced users, Boston, who provide a 'starter kit' and video guides, and San Francisco, who have an associated data academy, which provides free eLearning courses about data and the portal for public sector employees, could help inform an approach.

Providing active publisher support

Creating a comprehensive body of documentation and guidance does not replace active and direct support for publishers. Guidance takes many forms, some of it at scale, other ad-hoc and human.

Some of this support should be proactive, aiming to upskill data practitioners across London to publish, share, and use data better to make better decisions and enable innovation. A skills programme, building on the Data Skills Framework[25] could be attached to the Datastore to ensure quality and consistency, and foster creativity and impact in data access and use.

Building better publication tools into the platform to help publishers organise and prepare data for release should also be seen as part of providing proactive support. Building best practices into guidance can help embed good practice into day to day operations.

---

[24] Open Data Institute (2018), 'What data publishers need: synthesis of user research', https://docs.google.com/document/d/14vZJFUEJOkJEGOFTAPjRZ2FYxZfLQX3ct48K7bNsyl4/edit#heading=h.b3lnabharl6t

[25] Open Data Institute, 'Data Skills Framework' https://theodi.org/article/open-data-skills-framework/

Recommendation 5: Champion standards adoption and development

There are 33 local authority districts in Greater London, and many other agencies and organisations publishing or sharing data to the Datastore. It is unlikely that consistency of practice will emerge organically without some kind of coordination and convening.

There is no standardisation between boroughs and other agencies publishing to the platform, in terms of the data provided, formats, structure, or descriptions. This can make it hard to compare and combine data from different boroughs. As some interviewees and participants of the survey mentioned, providers of data to the Datastore often publish in other locations (e.g. their own websites, other portals), which can make it hard for users to know whether they have the most relevant data as they need to search multiple places.

The lack of consistency in the data affects the time users spend matching and standardising datasets. Participants at the user workshop suggested it would be helpful for the GLA to play a role in defining standards for data published on the Datastore, in order to help them integrate data from multiple sources for use in their analyses, products and services. This is related to Recommendation 3: Showcase data reuse.

The term "standards" can mean many things, and we are using it here with a broad definition, including the many activities and initiatives which can be undertaken to increase consistency and interoperability: from simple documentation of common practices (standards of quality), to the adoption of specific tools and formats (technical standards), all the way to the development of new open standards for data. The Open Standards for Data guidebook[26] provides guidance on many of those activities, with recommendations on how to find and adopt standards, all the way, when needed, to developing new ones.

In practice, we recommend the following first steps:
- Review the data currently published across GLA
- Identify areas for standardisation
- Start building a peer network of people to discuss and agree on standards

Recommendation 6: Encourage and facilitate collaboration

Many of the organisations surveyed through this project see the value in the Datastore as a discovery tool, rather than solely fulfilling a role as a technical platform. These organisations use the Datastore to discover which other organisations in London are sharing data and to identify potential collaborations. Some of the organisations who took part wanted to see the GLA help connecting different organisations who could be collaborators.

There are a number of organisations across London who are already supporting collaboration and innovation around data. The London Office for Technology and Innovation (LOTI) are working closely with London boroughs to run a variety of projects around data, such as reviewing borough approaches to the ethical use of AI and data, engaging with schools to raise student's aspirations in technology and improving the visibility of procurement activities across London authorities.

---

[26] Open Data Institute 'Open Standards for Data Guidebook', https://standards.theodi.org/

The Ministry for Housing, Communities and Local Government (MHCLG) are supporting local authorities, not just in London but across England, to improve digital skills and fund collaborative projects such as improving data standards for local community based services, through the Local Digital Fund. LocalGov Digital are aiming to support the visibility of local authority projects through their Pipeline in order to aid innovation and collaboration across the public sector.

Initiatives such as these show that not all the support and coordination of data publishing, sharing and use needs to be undertaken by the Datastore team. Encouraging collaboration – through convening, setting challenges to resolve specific problems, and resource sharing – is one way to rely on the community to take care of itself and create impact at scale.

Sharing beyond London
Fostering collaboration is the reasonable approach to dealing with many challenges common to the various agencies and boroughs of London. For the same reason, efforts should be made to collaborate closely with other city and city-region data initiatives to address common challenges.

In our engagement with users, we noted a user need for alignment between the London Datastore and other platforms, especially when using data to make decisions or solve problems at a scale broader than London itself. In one case, workshop participants expressed the desire for the London Datastore to enable discovery of data about neighbouring regions — which we understand to be the expression of the same user need.

The GLA are already thinking about the data infrastructure requirements for cities to be able to effectively share data, through projects like the Sharing Cities initiative. The London Datastore could be embedded into these types of projects, so that the sharing capabilities of the platform develop and align with other city data-sharing platforms.

The Desk Research Summary includes a number of examples of city data portal teams working collaboratively with other public sector organisations, industry or communities to establish new data sharing initiatives, or joining forces between similar cities.

Summary of Recommendations

Below is a recap of the themes and recommendations (1-6), with a reminder of the suggested paths to implementation.

Become a **better source of data**

(High priority, short to mid-term)

1. Improve the findability of the data
   a. Understand complaints about search and navigation
   b. Address concerns about "outdated" data
   c. Review categories for better navigation
   d. Provide curated guides
   e. Review metadata for a better experience and more effective search

2. Increase the variety and volume of data on the Datastore
   a. Engage with the community to prioritise which data to add
   b. Enable more technical means of access
   c. Continue the exploration of alternative access models

Invest in **a destination for insights**

(Slightly lower priority, short to mid-term)

3. Showcase data reuse
   a. Work with a community of data users
   b. Highlight insights, but not at the expense of data access

Become a **trusted guide and steward**

(High priority, short to longer term)

4. Document best practices
   a. Clarify what "good enough" looks like
   b. Provide guidance for data users
   c. Provide active publisher support

5. Champion standards adoption and development
   a. Review the data currently published across GLA
   b. Identify areas for standardisation
   c. Build a peer network of people to discuss and agree on standards

6. Encourage and facilitate collaboration
   a. Share beyond London

Appendix 1: Desk research summary

Approach

Desk research was conducted throughout the discovery phase. Early desk research helped to inform some of the structure for both the user and publisher workshops, as well as the interviews. Subsequent desk research helped to provide evidence that supports the findings from the user research, in order to inform this report's **recommended short term changes** to the Datastore, as well as **suggestions for longer term strategic plans.**

Aside from general research about the current London Datastore offering, the bulk of the desk research focused on these topic areas:

- Guidance around best practices for data portals and platforms
- Interesting examples of data portals and platforms beyond the London Datastore
- City level data strategies which support city data portals and platforms

Summary of findings

Guidance around best practices for data portals and platforms

Researching the landscape of existing guidance was a necessary first step towards ensuring that the proposed user research approach did not overlook any widely established recommendations.

As the London Datastore primarily functions as an open data portal[27], most of the relevant guidance for this discovery phase focuses on guidance around open data portals, rather than platforms. According to the Recommendations for open data portals: from setup to sustainability[28] report, the purpose of an open data portal is to:

- help people find the data they need,
- ensure that data accessed via the portal continues to be relevant, useful and usable,
- monitor and improve the quality and timeliness of data accessed via the portal,
- keep pace with data technologies and services, and user needs, as they evolve.

Open data portals are primarily concerned with helping users to access, use and share the data, so good open data portals must be designed to make it as easy as possible for users to engage with them. Consistent themes from various sources of guidance (An evaluation of U.S. municipal open data portals) (GovEx Labs: Open Data Portal Requirements) suggest that there are five different factors that portal owners should consider in order to enable users to get the best value from their portal:

---

[27] Leigh Dodds (2015) 'What is a Data Portal', https://blog.ldodds.com/2015/10/13/what-is-a-data-portal/

[28] European Data Portal (2017), 'Recommendations for open data portals: from setup to sustainability', https://www.europeandataportal.eu/sites/default/files/edp_s3wp4_sustainability_recommendations.pdf

- accessibility and availability of data
- trust in the data available
- guidance to help users understand how to use the portal and the data
- tools that help people to engage with and integrate data from the portal
- provision for users to participate and give feedback

Portal owners should also plan for portals to be sustainable, by making sure that they:

- establish good governance
- have a funding strategy
- build strong data infrastructure
- continue to maintain and improve operations
- capture metrics of use

All of the considerations that apply to open data portals are also relevant to different data sharing models; good data infrastructure requires people, processes and technology that can support the data assets, regardless of where that data exists on the Data Spectrum. However, while open data portals appear to be a fairly well defined model with significant guidance around set up, best practices and sustainability, the requirements for a good data sharing platform are harder to define. As the ODI's Mapping Data Access project outlines, the data sharing landscape is complex and finding an approach that suits the requirements of a specific situation can be difficult.

Self-assessment against recommendations from desk research

As part of the desk research work, the Datastore team evaluated their current operations against the recommendations of two significant reports on data portals: the European Data Portal project 'From Setup to Sustainability' and the US GovEx Labs' 'Open Data Portal Requirements'[29]. The self-assessment was mainly positive, and yielded the following insights:

- The Datastore does not have a business plan, documented governance structure or published funding information, but one of the goals of this project and its future phases is to create a clear plan for the long term sustainability of the Datastore.
- Standards for quality and metadata are not set. This is consistent with insights from the research activities, pointing to a need for the clear definition of standards and firm enforcement of quality baselines.
- The current platform does not enable searching for datasets by terms contained in the data. This is a common problem with the current generation of data platforms and portals, but the acceleration of development and investment in data search from the likes of Google is likely to make this a must-have feature in the near future.

---

[29] GovEx Labs (2015) 'Open Data Portal Requirements',
https://labs.centerforgov.org/open-data/portal-requirements/

- A number of recommended features are not, at this point, implemented on the platform. These include the ability for data users to download data in bulk, and the ability for platform administrators to track and analyse consumer feedback.

Other data portals and platforms beyond the London Datastore

A number of city and national open data portals around the world take unique approaches to improving user experience in key areas.

There are a few good examples of open data platforms with design features to **make the platforms more usable**.
- Plymouth's open data portal has a [separate insights section](#) with search capabilities, which can help users to easily find relevant visualisations and tools.
- The [San Francisco open data portal](#) gives users the option to receive a dataset alert, when new datasets that are relevant to them become available or when existing datasets are updated.

Some open data portals, particularly in the US, have **created guidance for users** which can help them to navigate the portal.
- The [New York open data portal](#) provides some simple guidance about how to use the platform.
- The [Vancouver open data portal](#) has a clear guide for new users of the portal, as well as an advanced guide for technical users.
- The [Boston open data portal](#) provides [a starter kit](#) which helps users to understand the terminology around data and to think about how they could use the data. The Boston team has also produced a number of how-to video guides which help users to navigate the portal.

Many open data portals include a dedicated section to **showcase innovative uses of data** from the portal.
- For national open data portals in the EU, such as [France's](#) portal, it is now common practice to include a showcase section.
- [Barcelona open data portal](#).
- [Boston](#) and [San Francisco portals](#).

A few open data portals have good **built in feedback mechanisms** and **actively work with users** to engage with the data.
- The [Amsterdam open data portal](#) has a prominent feedback section.
- [Data Plymouth](#) organise meetups and events for people who are interested in using the data that is published on the Plymouth open data portal.
- Paris has its own meetup group called Paris Open Innovation, although the group focuses more broadly on digital transformation in Paris, not just on the use of data from the portal.
- [Barcelona open data portal](#) uses [Decidim](#) to engage citizens in conversations around data.

Other cities are **working collaboratively with other public sector organisations, industry or communities** to establish new data sharing initiatives.

- The Copenhagen City Data Exchange project has created a data hub to try to establish a data marketplace where the public and private sectors can exchange datasets.
- In New York, the Mayor's Office of Data Analytics (MODA), the Department of Information Technology and Telecommunications (DOITT), and NYC Digital have worked together to establish a city-wide data sharing platform, DataBridge.
- Milton Keynes has developed MK Data Hub, the technical infrastructure for their MK:Smart project. The Data Hub has helped them to create a data marketplace, where open and shared datasets are stored and can be accessed under specific terms.
- Cities such as Malaga and Porto are using the Industrial Data Spaces infrastructure to share data with, and access data from, fleet operators, to help improve traffic flow in their respective cities.

GovLab's Data Collaboratives Explorer points to a multitude of other examples of data access approaches where public and private sector organisations are sharing data for public benefit.

City level data strategies which support city data portals and platforms

Across Europe, national open data portals have been widely implemented. As of the 2018 edition of European Data Portal's Open Data Maturity in Europe report, 26 of the 28 member states of the European Union have their own national open data portals, and 81% of these countries have dedicated open data policies which cover the next five years. For most of these countries, these policies are embedded as part of a larger digital or open government strategy.

In a number of these member states, one or more major cities have also created their own open data portals. The European Data Portal conducted a study in 2016 which examined Europe's top eight open data cities based on best practices: Amsterdam, Barcelona, Berlin, Copenhagen, Paris, Stockholm, Vienna, and London. All of these cities had created their own open data portals and strategies, but were also considering open data to be an integral part of their smart city strategies. A follow up study in 2017 examined seven more European cities that had established good practices: Dublin; Ghent; Florence; Helsinki; Thessaloniki; Lisbon and Vilnius. These cities produced open data strategies as well, and had a similar focus to those in the first study, with their focuses centring on increasing efficiencies across the city by improving connectivity, as well as being more transparent with city data.

The 2017 follow up study makes a number of recommendations for cities who are establishing open data initiatives:
- embed open data initiatives into wider smart city strategies;
- focus on meeting the demands of users by making data that users value available at a city level;
- overcome skills or resource gaps by collaborating with other cities;
- show the practical use of open data through use cases and describe the value;
- engage with the user community by running community events;
- build strong commitment from senior stakeholders who can support the open data initiative;

- coordinate on the topic of access to data at a national level with local and regional authorities to overcome cultural, technical, financial and capacity barriers.

Some cities within the same country are working together to establish better strategies around open data and smart cities. In Finland, the [6Aika](#) - also referred to as the Six City Strategy - is a collaborative effort between the cities of Helsinki, Espoo, Tampere, Vanta, Turkuu and Oulu to work on a number of projects relating to sustainable urban development, employment and competence. Projects must include at least two member cities and usually engage with a combination of residents, companies and research, development and innovation organisations. Learning from the projects is fed back to the wider collective, in order to improve competencies across all six cities.

Appendix 2: User and publishers workshops summary

Approach

This discovery phase involved three workshops in November 2019 to explore the needs of organisations and individuals that are currently, and could potentially, publish to or use the London Datastore to share and access data.

Two workshops were open to the public: one brought together a range of data users and decision makers and the second sought views from publishers and potential publishers to the London Datastore. The third explored the current use of the London Datastore by data and technology teams in London borough councils and included seeking input to the future vision.

Each workshop broadly explored two main topics:
1. Current or potential problems with using the London Datastore.
2. Identifying what would need to be in place to overcome these challenges to enable better stewardship and use of data across London.

Attendees

The two public workshops attracted 26 attendees from across the public and private sectors, including representatives from local authorities, not-for-profit organisations and small and medium sized businesses.

10 data and technology officers from Camden, Croydon, Greenwich, Hackney, Hounslow, Lambeth, Tower Hamlet and Waltham Forest councils came together for the final workshop organised by the London Office of Technology and Innovation (LOTI) in collaboration with the Open Data Institute.

Representatives from the Datastore team at the Greater London Authority were present at each workshop.

Summary of findings

Users workshop

- Improving the search function and findability of data published and shared is key for users, coupled with the ability to signpost and link to related data/ functions within and outside the Datastore.
- There could be a role for the Datastore in defining the standards for sharing data about London to improve consistency and compatibility.
- Many operational boundaries for public services (eg fire brigade/ police) extend outside of the political boundaries of the GLA. Balancing the ability to signpost, combine and access data on the periphery is important for some users and standardising the way data is provided could facilitate this.

Publishers workshop

- The majority of challenges stem from ways of working and basic data management issues rather than the functionality of the Datastore.
- There is a clear demand for a clear strategic direction, standards, processes and guidelines to aid consistency in approach and ensure publishers have the right skills and knowledge.

LOTI/London boroughs workshop

- The majority of data published to the Datastore is done so for statutory reasons, rather than a desire to proactively share.
- The biggest challenge is the lack of shared skills, standards and processes across the Boroughs, rather than a lack of a technical platform to share data.
- There is an opportunity for collaboration in two areas:
  - Consistency - common guidance, skills, standards and approaches.
  - To address data needs that are bigger than a single borough - the need for aggregation, common formats, common platforms.

Further information

A full write up of each workshop can be found via the links below.

1. User Workshop
2. Publisher Workshop
3. LOTI Workshop

Appendix 3: Expert interview summary

Approach

These semi-structured interviews with four key stakeholders aimed to tap into the existing knowledge of user and publisher needs by exploring how their needs evolved over the years and how London Datastore should change to address those needs.

Interviewees selected covered:

- **The public sector perspective**: an interview with two members of London Fire Brigade — as they're currently both publishers and users of the London Datastore.
- **The private sector perspective**: an interview with a member of Thames Water — as they are currently a user of the London Datastore; and another interview with Citymapper — as a user of transport data from TFL data.
- **The perspective of a similar data platform**: an interview with a member of TFL — as they host data and share it with the London Datastore and others.

Summary of findings

Publisher's perspective

Incentives, blockers and challenges when sharing data and expectations from portals such London Datastore to meet their needs

Motivations to publishing or share data

1. **Engagement publisher – user:** "the biggest value I've seen from publishing on the Datastore is that people who are using our data often get in touch".
2. **Collaboration, bring knowledge together – "**In terms of sharing data, I think that's just such a crucial component in terms of being able to bring together intelligence, not just from an organisation but also from other organisations".
3. **Contribute to quality information –** by being transparent and open, publishers are contributing to great information, and that's a driver for them.

Concerns when sharing data

1. **Security and usage of data –** The most-mentioned worry that holders of data have is around sharing in a secure way. This concern comes with the fear of data that is shared ends up being misused/misrepresented or the consequences when the commercial interests of a third party are involved. As an interviewee noted: " What happens when the objective of a public council authority and a commercial company don't necessarily align? What happens to open data strategy?"
2. **Reputation, trust –** Interviewees pointed out some of the consequences of data misuse, such as not being perceived as a trustworthy organisation.
3. **Duplicating shared data –** one of the push backs for publishing into the Datastore from those that hold data is their not seeing the value that it would bring. As an interviewee noted: " If there was a strong public need for it [we would] probably share [data]" or as another highlighted: "We would need to understand the purpose and what

is the need in terms of sharing that information [...] when we already have a platform where we are sharing that information".

Blockers and challenges when publishing to the London Datastore

1. **Format and quality of data** – noted as being limited
2. **Limited metadata view** – which makes difficult to give more context to the data
3. **Political issues** – as in collaboration being affected by the political climate
4. **Risk of "clunky" insights** – If the Datastore were to move towards an insight platform

User's perspective

Incentives, blockers and challenges when accessing London city data and expectations from portals such London Datastore to meet their needs

General user needs and frequent blockers

1. **Updated data** - to provide more context and significance to the data they want to use
2. **Navigate in a meaningful way** – to effectively find the data they require and decide whether is the appropriate one for their needs
3. **Get access to relevant data** – and know the source that contains such information

Trends on user needs and future challenges

1. **Users increasingly asking for more data** – as they experience the value of data
2. **Real-time data** – to generate insights that come from live data and act upon them
3. **Various user types and evolving needs** – when asking about the trends of uses of data, interviewees mentioned different user types that are emerging and what their needs and incentives to access data about London might be:
   a. Users learning data skills
   b. Users that need to give context to the data available
   c. Users accessing data to inform decision making
   d. User outside London that want to compare data with other city areas
   e. User advocating for transparency and openness

Appendix 4: Survey analysis summary

Approach

With the objective of understanding attitudes and needs from both users and non-users of the Datastore, we ran a survey for a month with general and filtered questions — some of the questions were only shown based on earlier answers.

The survey was circulated widely by the ODI and GLA teams, and promoted on the home page of the Datastore for several weeks. It got a total of 124 responses and covered questions in relation to:

- Behaviours when accessing or publishing data
- Motivation/incentive to use open city data
- Attitudes and expectations towards open/shared city data
- Demographics

List of survey questions

1. How often do you look for data about London?
2. What are the main reasons you'd look for information about London?
3. How do you typically access information about London?
4. How likely are you to look for information about London to answer questions…
5. How relevant are these areas to you? - either for professional or personal reasons.
6. Is there any other area that we haven't mentioned above?
7. Have you used data available from the London Datastore?
8. How useful has the London Datastore been to you so far?
9. From your perspective, what would make the London Datastore more useful?
10. How easy is it to access data (or share data, if you do so) through the London Datastore?
11. From your perspective, what would make the London Datastore easier to use?
12. Have you used any other platform that contains open city data?
    - if yes, please Specify: which one how have you used it (accessing the data or publishing/ sharing city data) how does it compare to the London Datastore - e.g. In terms of quality, quantity and/or updated data
13. What kind of application/uses do you foresee created with data about London, and who do you think should be creating that value?
14. Do you or your organisation hold (create; collect; use/maintain) any data about London?
15. What are your main concerns when publishing data or sharing it privately? - eg. Sharing process, security, quality, format, usefulness, privacy/ethics etc
16. Have you published or shared data using the London Datastore?
    If yes, where does Datastore not fully address these concerns/your needs when publishing/sharing data?
    If no, would you be interested in using Datastore to publish or privately share data?
17. How much do you trust the London Datastore to be the platform through which you could provide and manage access to such data?
18. Why did you choose that level of trust?
19. What is your gender? [*]

20. What is your ethnicity? [*]
21. What is the highest degree or level of school you have completed? [*]
22. What is your employment status? [*]
23. What type of organisation do you work for?

[*] Questions on demographics, and the choices given to survey participants were provided by GLA for consistency with their current practices.

Summary of findings

Ways of making the London Datastore more useful and easier to use

Those that are current users of the London Datastore rated it as an average of 7.8 useful and 7.0 easy to access the data. These are the themes that emerged when we asked about ways of improving its usefulness and usability and from the answers to a question asking for comparison with other portals – questions 9, 11 and 12 respectively :

1. **Up to date data**
2. **Broader datasets and types of data**
3. **Metadata and different views for insights**
4. **Improving navigation and search function -** including expanding categories and better description of them
5. **Offering signpost and updates indicators**
6. **Offering different formats** – and allowing interactive visualisation and charts, but also better integration between data and analytical outputs

Concerns and blockers for data holders

This group includes both current publishers of data in the London Datastore and those that hold city data, but not necessary shares it. We asked them an open question of concerns when sharing data in general –question 15– and another question where they ranked how do they trust the LDS to be the platform where they share the data they hold – question 17. These are the themes that emerged:

Concerns when sharing data

1. **Quality** – also reference to formats, accuracy and consistency of data
2. **Usefulness** – a worry about whether the data they'd share is useful enough
3. **Privacy/ethics and security**
4. **Usage** – including concerns on misinterpretation and/or misuse

Blockers when publishing on the London Datastore

Although a great number of stewards of city data trust the London Datastore to publish the data they hold, there were some reasons why others do not fully trust it:

1. **Lack of engagement** – and not fully understanding the London Datastore aims and plans
2. **Trust** – The London Datastore needs to prove that it is trustworthy by defining their values or the way data is managed (governance, curation, type of data they would make available, etc)

3. **Unclear curation process** – it's being perceived as not precise enough

Comparison between publishers, users and non-users of the London Datastore

	NON- USERS	USERS	DATA HOLDERS
Frequency in looking for information about London	2-3 times per year (26.1%) Once a month (26.1%)	Weekly (46.5%) Daily (18.6%)	Weekly (36.6%) Daily (24.4%)
Reasons to look for information about London	As part of my job (60.9%) Likely to answer questions about subset of London (47.8%)	As part of my job (76.7%) Highly likely to answer questions about subset of London (58.1%)	As part of my job (80.5%) Highly likely to answer questions about London as a whole (48.8%) and as a subset (48.8%)
Preferred way of accessing data	Ready-made insights (59.1%) Downloading data (50%)	Downloading data (88.4%) Ready-made insights (48.8%)	Downloading data (82.5%) Ready-made insights (50%)

Full survey report

The complete survey report - including graphs and anonymised answers, can be found here:
[2019 London Datastore Survey - Summary of all responses](#)