What Lies Beyond Open Data Portals

Exploring the potential of advanced data sharing for organizations to realize the value of their data

CHORDANT®

Introduction

Open Data Portals gained their traction in the last 10 years. They have been a great initial vehicle for cities to begin sharing their data and encourage transparency & innovation. However, this has not been without its challenges. Going back several years, everybody made a mad rush and they made all these wonderful data sets available, but nobody was coming to use them. Today, data stewards increasingly are looking beyond data availability, making actual data usage a key policy goal. Data should spark civic engagement and a free flow of new ideas. It should also break down the silos and foster more cross-departmental collaboration [1].

Cities often have no way of knowing who is using their open data or how valuable it is to those entities. Some Open Data Portal owners have resorted to running surveys to find out that information (see for example [2] and [3]). Even with the surveys, the information is only as fresh as the period covered by the survey. Cities will typically spend several hundred dollars in startup costs and several hundred more in annual expenses to set up Open Data Portals. There are also many unexpected costs. As discussed in [4], these hidden costs can include:

- costs from legacy system data that require reformatting
- quality-related costs to keeping open data fresh and up-to-date
- legal costs to comply with open data legislation
- liability costs in case something goes wrong, such as publication of nonpublic information
- public relations costs that can occur when a jurisdiction generates bad press from open data.

There are many other views related to open data [5], some of which include:

- revenue earned by publishing data can be used to cover costs of generating and/or disseminating it or to fund other city and non-profit activities
- privacy concerns may require that access to data is limited to specific users or to sub-sets of the data
- collecting, cleaning, managing, and disseminating data are typically labor- and/or cost-intensive processes – whoever provides these services should receive fair remuneration for providing those services
- sponsors do not get full value unless their data is used appropriately – sometimes this requires quality management, dissemination and branding efforts that can best be achieved by charging fees to users
- targeted end-users cannot use the data without additional processing (analysis, apps etc.) – if anyone has access to the data, none may have an incentive to invest in the processing required to make data useful (typical examples include biological, medical, and environmental data).

Although a great first step, Open Data Portals are often not enough to satisfy all the data sharing requirements for cities and their ecosystems. Discussed below are some of the key functions necessary for the evolution of Open Data Portals that can help solve above outlined challenges and improve both the data intrinsic value and the overall return on investment.
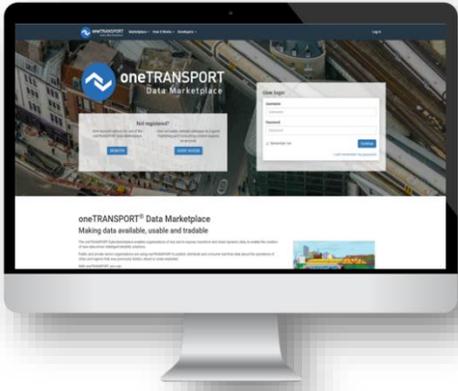
Sources
[1] https://www.govtech.com/data/Are-Open-Data-Efforts-Working.html
[2] https://www.ontario.ca/form/survey-open-data-catalogue
[3] https://www.smartcitiesdive.com/news/philadelphia-open-data-portal-survey-community-engagement
[4] https://www.governing.com/columns/tech-talk/gov-open-data-cost-problems.html
[5] https://en.wikipedia.org/wiki/Open_data

CHORDANT®

User Registration
*Be aware of who is using your data and how*

Open Data Portals by nature provide open data, meaning that users can freely explore and download of data, while cities have no knowledge when it happens. This allows for an easy access, but cities could choose additional access options. For example, they could ask users to fill in their information to gain access (open registration) or have users request access (closed registration). This allows cities visibility into who is consuming their data. By knowing their data consumers, cities can better target the available data, send periodic updates on new data availability and engage their users in more meaningful ways.
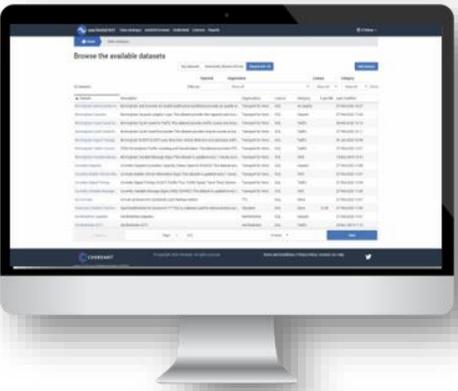
Multiple Data Sharing Options
*Be confident that your data is not getting into the wrong hands*

Cities usually share data for transparency purposes. In some cases, they are mandated to do so by law. Nevertheless, cities also have data that may be confidential or containing personally identifiable information (PII). This type of data cannot be shared in an open fashion. Such data may be restricted to sharing internally between city departments or very selectively with outside parties. In this case, having options to share data privately (within organization) or selectively (with approved parties only) allows cities to share more of their data. This in turn helps break down siloes and enables cities' ecosystems to create new and valuable data-driven services.
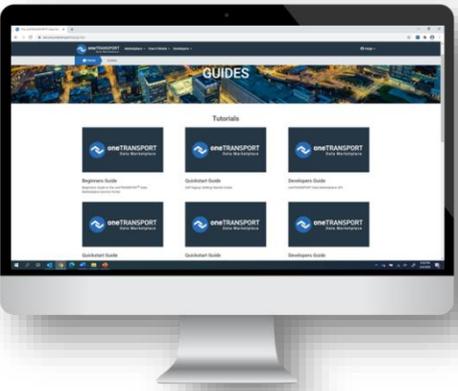
Ecosystem Engagement
*Be inclusive by leveraging a scalable sharing framework*

Open Data Portals often contain data from one city only. If a city wanted to invite their ecosystem (e.g. vendors, contractors, SME's, and other 3rd parties) to share their data with the city, Open Data Portals could not be used for that purpose. First, this is because Open Data Portals are open and do not protect confidentiality of data. Second, Open Data Portals only offer data consumption, but generally not the ability for data to be added. Third, Open Data Portals have only one 'Open Data License' that governs the use of all data. Creating data sharing agreements between parties is often a very daunting endeavor. Advanced data sharing allows for a creation of multiple data agreements that allow flexibility for both cities and 3rd parties.

CHORDANT®
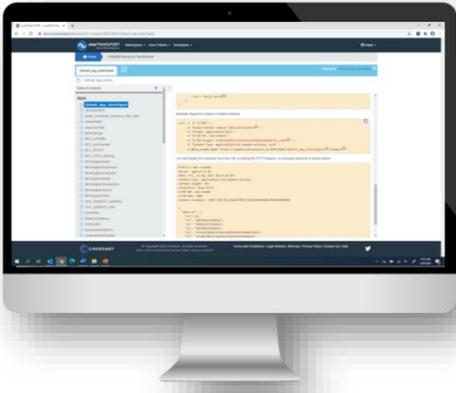
Real-Time Data Support
Be assured that data is up-to-date and of high-quality

Some Open Data Portals offer the ability to consume historical data via APIs, a scalable approach which makes it easier to directly integrate data into applications and services. However, most Open Data Portals do not offer real-time data sharing. Advanced real-time sharing platforms allow data streams to be consumed directly from the real-time systems. This means that there is no data replication in the data sharing platform, ensuring the security and privacy. Advanced real-time sharing functionality does allow real-time data to be archived to be used for later analysis, when needed. Data is also screened for availability and quality.
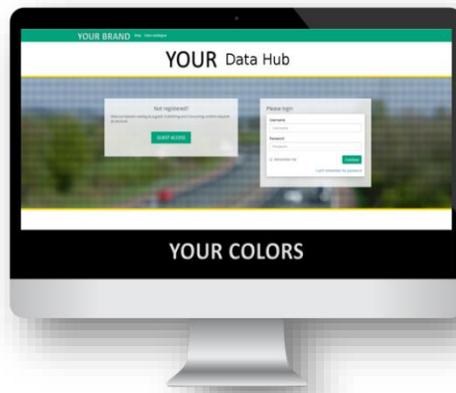
Collaboration Spaces
Be the leader in sparking civic and public-private engagements

Open Data Portals serve to expose cities' data, but they are not collaboration spaces. Advanced data sharing functionality allows cities to create multiple 'Data Hubs' that operate from the same platform backend. These instances provide temporary or permanent ways for a collaborating group to share data. For example, if a city were running a testing or a development project, it could enable data sharing between the entities that are engaged on that project only. The front-end can be customized to have a look-and-feel and the functionality specifically required for that project. Great examples include a city enabling V2X communication over 5G systems or testing of Connected and Autonomous Vehicles.

Data Scientists Platform
Be able to transform and present your data in usable forms

Most of the existing Open Data Portals offer historical data in the form of files or time-series that can sometimes be visualized directly in the portal. However, Open Data Portals do not offer any capabilities to transform, combine or process data. These functions can be very useful when data is being shared between systems and users that require data in different formats. This can mean rearranging the same data to present in different formats or normalizing different data to appear in the same data format. These capabilities are useful for both cities that do not have their own data science teams and those that want to focus on enabling solutions, rather than processing data.

CHORDANT®

Data Monetization

Be able to recover data processing costs & fund new initiatives

Although data monetization may not be applicable to all cities, some cities may choose to charge for some data such as Requests for Public Records. Cities may choose to monetize their data sharing in other ways. For a medium-large size city, the potential direct benefit is estimated at $15M with additional $60M in indirect benefits in 5 years that result from: city's own data sharing/sale, 3rd party data sale/resale, 3rd party app sale, neighbouring county access to the platform, and indirect benefits (reduced congestion, pollution,…).

Chordant Data Sharing Solutions

Chordant Symphonia™ data sharing product offers all these advanced features that can be enabled over time as the city's needs evolve. A city can leverage Chordant's existing data sharing service www.onetransport.io, powered by Symphonia. This existing service can be beneficial as a starter solution for many cities. It allows quick time to market and cost-effective implementation due to the economies of scale of a shared multi-tenant solution. Different Symphonia options are available when the city wants to establish its own dedicated place to share data.

Chordant data solutions include the following advanced functions:

- **User Registration** allows cities to be aware of who is using your data and how, so they can engage more users and improve their data offers over time
- **Multiple Data Sharing Options** allow cities to enable organizational sharing for departmental collaboration, selective sharing with the external ecosystem, and open data sharing with the community
- **Ecosystem Engagement** allows cities to be inclusive of their ecosystems allowing others to share their data and leveraging the extensive legal framework that provides custom data sharing agreements
- **Real-Time Data Sharing** allows cities to share time-sensitive data, but also to keep the data fresh, up-to-date, and monitored for availability and accuracy
- **Collaboration Spaces** allow cities to create collaborative spaces that allow selective data sharing relevant to the specific initiative and open to the initiative participants only
- **Data Scientists Platform** allows cities to transform data and create usable formats that are ready to be consumed by data scientists and application developers
- **Data Monetization** allows cities to monetize data to recover their data processing costs, fund new community initiatives and foster economic development in their regions.

Chordant and its partners can help cities identify relevant data sources and data providers for their applications. We can also help with understanding how data should be presented and relevant data sharing agreements. There exist different means to ensure privacy in Chordant's solutions such as abiding by the local privacy laws, screening the data to share, access controls, and other platform security mechanisms. Data can be stored locally (cloud or on-premise) and geographically restricted.

About Chordant

Chordant provides data sharing solutions that enable our customers to realize value from data. Our solutions power data-driven applications for Cities & Regions, Connected & Autonomous Mobility and Complex Infrastructure & Facilities. Chordant has been recognized by numerous analyst firms and organizations for its industry-leading solutions.

www.chordant.io

If you are interested in any of the functions above, have an interesting new use case idea, and would like to collaborate with us: chordantsales@chordant.io