Case Study:

DataDriven / NOLAlytics

Type: Website / Team

Organisation(s): City of New Orleans, the Sunglight Foundation, Code for America

Tags: open data, process, standards, website

DataDriven is the city data programme in the City of New Orleans, USA with four key aims:

1. Building an inventory of all departmental datasets
2. Building a citizen community to find new ways of using data
3. Publishing useful datasets
4. Promoting innovative work done with city data

The programme encompasses several different teams. The Enterprise Information Team within the Office of Information Technology and Innovation (OITI) manages the DataDriven website, the data inventory, and data-based web applications. Meanwhile, the GIS Team, also part of OITI, runs the GIS data catalogue, feed GIS-based web applications.

Another component of DataDriven is NOLAlytics, which is the analytics team within the Office of Performance and Accountability. This team aims to find data-driven solutions to the city's problems and partners with other city departments on projects which seek to make use of data in new ways. In many ways, this is similar to NYC Analytics, and Offices of Data Analytics set up across the UK.

Data Coordinators in each of the city departments make up the final part of this data architecture. Each of these officers is responsible for maintaining, tracking, and scoring the datasets within their own purview.

Background

The open data agenda in New Orleans emerged in the context of the recovery from Hurricane Katrina in 2005. Data such as population counts and numbers of blighted properties had been crucial to the recovery efforts. The NGO The Data Centre (formerly the Greater New Orleans Community Data Centre) was crucial in producing reports that informed these efforts.

This experience informed the decision in 2011 to establish New Orleans' first open data platform, data.nola.gov, which has now been superseded by datadriven.nola.gov. This portal gradually expanded as residents submitted requests for more information that they wanted to be shared.

In 2012, the City of New Orleans partnered with Code for America to build an app called BlightStatus, which is now inactive. This would help residents to track properties still in disrepair. This achievement fed into subsequent work through 2013 to publish further datasets, such as city boundary layers and authorised short-term rentals. The team also produced further web applications, such as 'Where Y'at', which provides a wealth neighbourhood information on each zip code in the city. These web applications are now an integral part of DataDriven's offer, providing added value to the raw data.

The first experimental dataset that was published provided information on curb lines, which in 2013 was incorporated into a GIS map using aerial imagery of the city.

In 2015, in response to the perceived need for a more systematic approach, the City of New Orleans formed a partnership with The Sunlight Foundation on the What Works Initiative, which is a Bloomberg Philanthropies project. This cooperation helped the city establish key principles that fed into a data policy. This also helped the DataDriven team understand how to make use of data to perform low-cost evaluations of the city's problems. The NOLAlytics team is responsible or this analytics work and ensures that DataDriven's work goes beyond providing an open data store.

In 2018, the DataDriven team established a comprehensive data catalogue, which drew together all of the open datasets held about the city, making it easier for data users to find what they are looking for.

The DataDriven website is hosted and managed by Socrata, which also provides bulk download and API services. However, DataDriven's GIS platform City of New Orleans GIS is on a separate webpage and is powered by ArcGIS.

Important considerations

Data publication and categorisation

The data publication process begins with a request submitted by a 'data champion', who can be anyone interested in a particular dataset being published. The data champion fills out an inventory questionnaire and then scores the dataset using a scorecard, based on features including priority and value, technical challenge, data classification, and data quality concerns. Each dataset is then passed on to a data steward within the relevant department who is responsible for adding metadata, and curating and updating the dataset.

The data classification meanwhile is used to determine the suitability of a dataset for public release. This involves assigning the dataset to one of these categories:

1. Public data, that can be publicly disseminated
2. Protected data, that can only be shared internally because of regulative protection or because it includes individually identifiable information
3. Sensitive data, that poses security concerns and could be misused

Real-world application

The NOLAlytics team views the presentation and visualisation of data as only the first step in a longer process. The team has a strong focus on actively finding new ways to use data to improve city governance. They apply a Use Case Typology to classify the different ways that data can improve city administration.

1. **Finding the needle in a haystack:** This includes identifying specific opportunities to create impact, using tools such as predictive modelling. An example is predicting where health and safety violations will occur to better target a small number of inspections and establishing regular audits of identified high-risk businesses.
2. **Prioritising work for impact:** This involves using data to prioritise work based on potential impact or complexity instead of when the request was made.
3. **Early warning tools:** This involves flagging areas where early or proactive investment is required, e.g. predicting homes where lead poisoning is most likely.
4. **Better, quicker decisions:** This involves using real-time information to improve operational decision-making.

5. **Optimising resource allocation:** This involves drawing on data to ensure that scarce resources are used in the best possible way.
6. **Experimenting with what works:** This involves trying out different approaches and determining which is most effective, such as testing the effectiveness of SMS vs email notifications.

All of these approaches require well-organised, standardised, complete datasets, which is the role and aim of DataDriven.

By applying their Use Case Typology and learning from best practice observed in other cities across the USA, the NOLAlytics team has worked on a range of challenge-based projects across the city administration.

For example, NOLAlytics ran an Ambulance Posting Location Optimisation project in 2017. This involved developing an algorithm that ranked locations around New Orleans by impact, based on data on traffic patterns, 911 calls, and ambulance availability. By using locations suggested by this algorithm, the team achieved improved response times in those areas with historically low response times. The openness and consistency of these datasets helped to ensure the success of this project.

Content and quality

The DataDriven portal includes an Open Data tab that allows users to search for datasets, as well as displaying 'featured datasets' and 'new datasets'. The featured datasets are 'Calls for Service', which relates to calls made to the police and '311 Service Requests'. These indicate a continued strategic emphasis on public service transparency within the Enterprise Information Team that manages this portal. The search function works well and links to the right datasets even when typing in search terms that are not included in the dataset title or description.

As an alternative to this free text search, there is a full data catalogue, which allows users to explore the data using filters.

Overall, there are 263 datasets, many of which are expansive, apply to long time periods, and have thousands of rows of data. Datasets are well consolidated and there are no duplications of equivalent datasets. The catalogue includes a wide range of datasets available to data users across all areas of city life, from demographic and economic data, to city administration metrics and land ownership data.

Datasets are also filterable in multiple ways, including by category, dataset type, city department, and by tags that have been attached to them. They are also generally up to date, with some datasets, such as those on building permits, updated daily.

Many of these data have significant commercial potential. For instance, the 2018 Market Value Analysis Report is openly available for download and helps those working in both real estate government to identify trends and patterns. Another interesting resource is a map with the location of all businesses which provide 'preservation services', such as repairs, painting, and plumbing. This is a very useful dataset for residents and businesses as a city prone to regular flooding.

Functionality for data users

There are several different ways data users can work with and manipulate statistical datasets. As with many other data stores, they can link to the Socrata API and export the data into a format of their choice (for example CSV, RDF, RSS, TSV, or XML).

However, there are also additional functionalities, including the ability to create bespoke visualisations of the data, selecting columns and rows, and choosing their preferred type of chart. Data users who have set up an account on DataDriven are also able to save visualisations to their account for future reference. They can also share datasets on social media and contact the Dataset Owner directly using an online form.

Another component of the DataDriven portal is the data products that allow the user to interact with the data in more detail. These include a web app called 'Streetwise' that tracks a live feed to show all current traffic or flooding incidents in the city and allows residents to directly report any incidents that they sight. Users can view the exact location each incident occurred and when the incident was reported.

Another product, 'Sanitation Services', allows users to select an address and view the rubbish collection schedules for each type of waste as well as the provider responsible.

DataDriven does not provide purely operational data. ResultsNOLA, developed by the NOLAlytics team, provides data showing to what extent the City of New Orleans meets the strategic priorities it has set itself. This includes all of the latest figures on measures that relate to each strategic priority, from the homicide rate to the number of smoke alarm installations and the percentage of streetlights functioning. This can help to provide both strategic information, and put operational data in context for data users.

Public outreach

A focus on public engagement has been built into the DataDriven plan. Among other responsibilities, the Enterprise Information team engages with the public about data. They do so through the local brigade of Code for America, which aims to strengthen links between technology and government, and through neighbourhood meetings. The public is also able to interact easily with the city authorities using the site, for instance by directly pitching an analytics project idea to the NOLAlytics team or by getting in touch with Data Owners for each dataset on the site.

Blockers and challenges

DataDriven's oraganisational structure and data publication system is complex and systematic and works well when all the correct pieces are in place. However, to function properly, the system relies on a Data Coordinator being assigned in each department. When there are political changes such as the election of a new mayor, this system can be disrupted.

Also, whilst New Orleans was an early adopter of open data, in the first few years the city's work in this field was fairly haphazard and was not backed up by a clear strategy. It was only the articulation of a clear policy framework in 2016 that enabled the city to have such a large and comprehensive open data and analytics programme.

What can Greater Manchester take from this?

- It is important to note that New Orleans functions in a very different political and administrative environment to Greater Manchester. Nevertheless, they have been able to structure their teams and services around data ownership, stewardship and release.
- Assigning a Data Coordinator to each local authority service could help ensure clear and continuous lines of communication between the team running the local data programme and council officers on the ground.

- Recording who exactly is the contact point for each dataset also allows data users to easily reach the person responsible in case of any queries or concerns about a dataset.
- As long as there are clearly defined areas of responsibility, it is possible to run an open data and analytics programme that encompasses multiple teams which specialise in different aspects of the programme, like user interface, GIS systems and applying data to policy.
- Whilst it is important to take data publication seriously, this should not be seen as the end of the process. There should be just as much focus on making the best use of what is already available as there is on publishing new datasets.
- A framework like the Active Use Typology could help focus minds and ensure that the identification of possible new use cases is always at the heart of Greater Manchester's open data strategy.
- A dashboard with a comprehensive overview of city metrics similar to ResultsNOLA would provide residents and local government employees with an easy understanding of how Greater Manchester is performing and what areas need improvement.
- Drawing up a complete inventory of all datasets that relate to Greater Manchester would help ensure that the longer process of data publication is guided by evidence and a clear, systematic approach. It is easier to decide on a publication programme if it clear from the start where opportunities lie.

Find out more:

https://datadriven.nola.gov/nolalytics/

https://www.bloomberg.com/news/articles/2015-04-22/how-new-orleans-louisville-and-san-francisco-are-using-open-data-to-solve-problems

https://datasmart.ash.harvard.edu/news/article/using-predictive-analytics-to-combat-rodents-in-chicago-271

https://journals.sagepub.com/doi/10.1177/1098611117709785