Navigating Open Data Standards: Lessons from the Citadel Project

1. Citadel on the Move: Building a Data City

Citadel on the Move set out with a simple vision 3 years ago: To make Open Data an achievable reality for every city in Europe. The project, a 4M Euro-funded EC Smart City flagship[1], has since helped over 120 local government organisations around the world open their data.

To support our extensive global outreach campaign, **Citadel** developed a range of new open source tools - data convertor and an app generator - that make it easier for all governments, especially the small ones that often get left behind, to Open Data *and* use their data to unlock smart city innovation.[2]

During the course of our work with more than 120 associate cities, **Citadel** realised the project was not only helping to unlock local government data on an unprecedented scale, but also uncovering a first-of-its-kind evidence base of findings about the realities of Open Data at the local level. From level of experience through to the most popular data formats, **Citadel on the Move** has revealed a unique snapshot about global open data trends at the local level.

Citadel gathered first-hand data and experiences from 63 countries worldwide, including 46 here in Europe. This wealth of information has allowed the project to develop a unique understanding of what it takes to start building a 'Data City' – A city that truly embraces the potential of Open Data. From navigating the minefield of data standards to making your local area a magnet for data-driven businesses, **Citadel** has drawn on our extensive experiences to create a simple, clear guide of business recommendations for Local Authorities everywhere to open their data.

The following pages are set out as your end-to-end companion for building a Data City. Starting with the Business Case for action, through the best way to open data to stimulating data-driven business in your local area, this document gives you everything you need to start building your Open Data City.

---

[1] Citadel on the Move was funded under the EU's 'Competitiveness & Innovation programme in 2011. For full details of the funding please visit: *http://ec.europa.eu/information_society/apps/projects/factsheet/index.cfm?project_ref=297188*
[2] *www.citadelonthemove.eu*

2. Your Business Case for Opening Data

Creating new services, managing resource use, better decision making and evidence-driven policy formulation - cities across the world are discovering the power of Open Data to help solve civic problems and create economic opportunities!

Data is government's newest resource. Collecting and sitting on vast quantities of it for many years, the Public Sector has only just started to discover the potential of utilising this raw material. Never before have cities been in a position to generate self-digital-awareness by combining their technology infrastructures with their own data. From making evidence-based decisions to holding themselves to account, government at all levels has the potential to transform their operations to be more effective and efficient.

However, its when public data – transportation, health care, education etc. - is opened up outside of government that the possibilities for innovation become limitless. From fueling the formation of new city services that tackle urban problems to the creation of new markets, businesses and jobs, Open Data enables entrepreneurs, developers, and businesses to spark new economic development within cities even during the toughest of fiscal times.

Data is truly a great leveller. All cities and administrations have the same potential to harness the wealth of benefits heralded by the Open Data movement. Yet **Citadel on the Move** understands, from experience working with more than 120 cities across the globe, that opening public data is often easier said than done. Whilst technology challenges can be easy to address with help from the Open Community, the largest barrier to be overcome is convincing departments and organisations to release their data in the first place.

Data 'huggers' (those who are reluctant to publish) are often overly concerned with the negative aspects of opening data, with fears ranging from maintainance issues and privacy through to worries about quality. **Citadel on the Move** however, has demonstrated that the benefits of opening data far exceeds, and often directly contradicts any perceived negativities, and as a result has been able to formulate three key drivers, that form a compelling Business Case for cities to go 'open'!

1. Achieve Significant Cost Savings

Opening up public data has significant potential for cost savings and improving service efficiencies. All cities and administrations spend significant amounts of time and resource in answering citizen queries arising from policy formulation, or from local news and events. Researching the answer for these queries ultimately results in giving the questioner being given access to a piece of public information. Opening up data sets and publishing them online with easily searchable interface reduces the direct impact and cost of servicing these requests.

Additionally, openness helps cut down unnecessary public expenditure, as Public Bodies are made more accountable for financial discrepancies. A 2013 CapGemini report on the Open Economy[3] found that in California, USA, the state transparency portal (that cost $21 thousand to implement) saved the state over $20 million when visitors to the site identified unnecessary expenditure. The savings were generated when users saw an audit that showed many of the vehicles in the states fleet were not needed.

Case Story: Freedom of Information in Manchester
Manchester City Council (one of Citadel's founding partners) estimates that opening up its data saves them around £8.5 million. Previously this money would have been spent through resource time used

---

[3] http://www.capgemini-consulting.com/resource-file-access/resource/pdf/opendata_pov_6feb.pdf

searching for data to answer public 'Freedom of Information' requests – a service that must be provided by Councils by law. Now requesters are directed to the cities open data portal to find information themselves which resulted in a reduction of over 1,000 official requests each year.

2. Unlock the Value of Your Greatest Asset – Your Constituents.

Enabling businesses to tap into city data to provide their customers with useful information and insights stimulates new business and economic growth. Open Data can help app developers, urban planners, and others understand a city's problems and manage city services in ways that improve the quality of life and business prospects for its residents. Organizations such as CityMapper[4] have taken open transport data including bus and rail timetables along with real-time feeds to create optimal routing apps for smart phones to help people navigate across many of the worlds complex cities faster and easier than ever before.

Even people who aren't building apps can help Cities enhance the quality of their services. In 2011 UK Government opened its transport node data and found many inaccuracies, some bus stops were listed as being in the middle of lakes! Updating the dataset through a public procurement would have been an expensive process, but by making the data open, a community of volunteers cleaned it for public use in less than 3 months removing all data inaccuracies.

Case Story: New York City Big Apps[5]
New York City's BigApps contest taps into the brightest talent in the city to bring people together to work on and solve major social challenges affecting the cities residents, visitors and businesses. The programme encourages multidisciplinary teams of programmers, engineers, designers, marketers, civic professionals and entrepreneurs to work together to create new products and services. Incentives are provided in the form of a pool of mentors from innovative tech companies such as Google, Microsoft, and Facebook with over $100k in cash prizes. To date, the competition enables New York to return $10 on every $1 invested in the competition making it a cost-effective way to tap in to technology and data to deliver innovation and quickly spin-off new business start-ups.

3. Meet Transparency and Accountability Laws

New National directives are starting to stipulate that Local Government must increase the transparency of their operations through enhanced visibility of financial spending, project budgets, effectiveness etc., all by publishing open data. A timely example of these new laws in action comes from the UK who has just released its Local Government Transparency Code 2014 mandate, complete with statutory prescription and suggested datasets for release to fixed timetables. Other countries will be soon following suit.

Case Story: Open Budgeting in Helsinki[6]
In 2011 budget chiefs in the Helsinki area first opened its financial data, along with an application called OpenSpending. Through simple visuals the application helped constituents concretely understand how much budget is needed annually by different elements of the municipality, from general areas like health care through to specific surgeries and individual community centers. The data visualizations helped the community better understand government processes and enabled enhanced feedback and discussions about budget allocations with city management. As a result one of the key benefits from opening up public data was increased civic participation and democracy leading to improved relationships between people and government.

---

[4] http://citymapper.com
[5] http://nycbigapps.com
[6] https://openspending.org/city_of_helsinki_expenditure_2009_2010

The rest of this guide takes you through all the steps you need to consider when developing your own business case for your own Data City strategy. From which data sets to publish first, through to who needs to be engaged and what kinds of business models exist for helping your communities release value from your data, **Citadel on the Move** shares with you the lessons and recommendations generated by three years of on-the-ground working with cities across Europe to help you **stop worrying and start loving your data!**

3. Opening Your Data: Doing it 'Right' from the Start

Citadel knows from first hand experience that opening up government data is often easier said than done. What's more, many of our Associates have found that if you make the wrong decisions when publishing your data, it can be costly and time consuming to fix them later.

This challenge is why **Citadel** places an emphasis on doing things 'right' from the start. The information below lets you make smart choices about Open Data that won't leave you pulling your hair out in the future.

When to Open?

Citadel recommends you to **open your data right now**. Too many Local Authorities focus on getting everything perfect before opening their data. This concern for perfection in practice means most datasets will never be published because the amount of work required to get them 'perfect' is more trouble than it is worth. Bear in mind that **Open Data leads to better data**. You can always improve your datasets after they have been opened but in the meantime it is much better to allow your creative community to be using that data to make your city a better place to live.

How to Open?

Citadel recommends you to follow these 5 common-sense practices in all your Open Data work:

1. Use Standard Terminologies – Make sure the same term is always referred to in the same way. For example, in address data, you will often see some entries spelt 'Street' while others are spelled 'St.' Different spellings make it difficult to search your data properly and mean fewer people will use it. To solve the problem, agree on a set of standard terminologies (spellings) and make sure your staff know what they are when they create or check data.

2. Fix Blank Data Cells – Make sure every 'cell' in your dataset is filled in. One of the biggest problems with datasets is missing information that makes it difficult to use. To solve this problem, make sure your staff goes the extra mile to fill in blank entries and it will pay dividends when re-using the data.

3. Understand Your Data's Back Story – Make sure you understand for what purpose your data was first collected. Every dataset was collected for a specific purpose, be it a financial record keeping or the delivery of a specific service but most local authorities don't keep a record of this information. To solve this problem, wherever possible make sure your staff identifies the reason a dataset was collected and where it came from so they can make better judgements about what to do with it.

4. Provide Maximum Information Available – Make sure you include all available information in the dataset. Often Open Data contains only the basic information even when the local authority holds more details for each data point. By adding every piece of information you have, you will make it more likely that creative people will reuse your data. The only exception to this is names, addresses or other identifying details classed at sensitive personal data. Also provide information about how and when the dataset was collected including the author, description and update policy.

5. Update Regularly – Nothing makes a dataset more likely to be reused than knowing you can trust that information. Make sure you check your datasets regularly (wherever

possible) to keep them up to date and check for consistency and quality. How often is up to you, it could be every week or every three months but people need to know.

What to Open?

Citadel understands that deciding which datasets should be opened first is daunting. The information provides guidance on how best to get started:

Which Datasets to Start With – Local government has access to many different kinds of data. Knowing what to focus on opening is important to make sure you don't waste time and resources releasing information that people don't want. The initial priority should include datasets for your city that have proved valuable elsewhere or that reflect known local interests. If transparency is key then financial datasets may be your starting point– although it is important to remember that these sets are often the least likely to be reused. If you wish to stimulate innovation and new service creation then Point-of-Interest datasets of more general interest such as those related to tourism, crime, public health and transport may be more useful

What Information to Include – A good point-of-interest dataset will always contain certain pieces of information any re-user will need. Make sure you have the following information each under its own column:

- **I.D. –** Number value for the entry e.g. 17
- **Title –** Title of the entry e.g. Gallery of Modern Art
- **Category –** What type of entry it is e.g. Art Museum
- **Description –** A description of the entry e.g. A museum specialising in art from the period 1850-1990
- **Address –** The address details of the entry e.g. Royal Exchange Square, Glasgow
- **Postal Code –** Postal reference code for the address e.g. G1 3AH
- **Location –** Latitude and Longitude value for the location of the entry in the form: Latitude[space]Longitude e.g. 55.860021 -4.252045

The above are *basic* requirements for a good dataset. You should also include any other information you have available in separate columns (examples for an art gallery might include opening hours, disabled access, prices)

What File Format to Use – Once you have a well-structured dataset, you need to save it in a specific file format. Formats are important because they determine how the information can be viewed and reused by creative people. At **Citadel** we recommend using .CSV (Comma-Separated Values). The reason is simple: CSV can be saved and opened from any text editor, and also an Excel file and gives the most flexibility for creative people to reuse your Open Data. For more advanced releases, we also recommend using .JSON (Java Script Object Notation) because this format is very popular with app designers. **Citadel**'s convertor[7] can transform your CSV files into JSON for free. Remember to make sure anything you publish in JSON you *also* publish in CSV.

More information about formats and overcoming common challenges can be found in Chapter 5.

Where to Open?

---

[7] http://www.citadelonthemove.eu/en-us/opendata/convertmydataset.aspx

Once you have created a well-structured, useful dataset, you need to determine the best place to put the data and under what conditions i.e. will you allow unrestricted reuse or do you need to place certain restrictions.

Open Data Portals: The place where you make your Open Data accessible for people to download and use is known as an Open Data portal. Choosing the right portal means considering three key questions: 1) Where will the maximum number of people see my data? 2) Where will it be easiest for me to update my data? And 3) How easy is it to access, preview and reuse my data?

Citadel recommends using any publicly available Open Data portal to publish datasets rather than just a standard page on the council website. Council websites can be used as *additional* storage locations but should always be used in combination with a website dedicated to publishing Open Data.

In many countries, the National government provides a portal that allows local authorities to publish without the cost of having to create their own site. Good examples of these national portals include data.gov.uk, data.gouv.fr and dati.gov.it. For some more advanced local authorities, you may wish to follow the example of Leeds in the UK (leedsdatamill.org) or Ghent in Belgium (data.gent.be) and create a dedicated portal for your area.

Citadel recommends that all Open Data should be published in the form of a CSV file that can be downloaded. Wherever possible, it is also recommended that the data be made available both as a downloadable JSON file and via an API.

Open Data Licencing: An Open Data licence is an agreement that regulates how and when people can use the data they download from you.   The licence tells people about what conditions your data can be used under (e.g. can it be used for commercial purposes or not?) and what things they have to do when they use the data (e.g. do they have to quote you as the source or not?). While a wide variety of licenses exist that qualify as 'Open Data', many place restrictions on the user about how they can use the data.

Citadel recommends using the 'Creative Commons' family of licenses (creativecommons.org) because they are very popular with Open Data providers, widely accepted and simple to understand. **Citadel** specifically recommends using CC-0 'No rights reserved' wherever possible. For data where your stakeholders insist on being credited, **Citadel** recommends Creative Commons Attribution 4.0 (CC-BY) which preserves attribution of the dataset to the owner. Citadel cautions against public institutions writing their own licenses as in the past local and national governments have made errors in the past

Who Can Open?

Identifying the Right People: You will need to identify the people who own your data and get them onside for your programme to succeed. **Citadel** recommends identifying the Data Manager or Service Manager from each department you are targeting. Once you have these names, focus on getting them together in a room for a meeting. Make sure you clarify that Open Data at your authority will be run as a central programme and let

them understand the benefits of taking part. Then ask each of them to bring you their datasets so you can decide what to open and in what order

4. Promoting Your Data: Unlocking Innovation Potential

"It's about turning government into a platform for open innovation. Data by itself is useless. I can't feed my baby daughter data, as much as I'd love to because I love data. It's only useful if you apply it to create an actual public benefit. You need appliers — you need entrepreneurs to know data's there available in order for them to turn it into awesomeness." – **Todd Park, CTO, United States**

Opening data is more than just about publishing it. Your data has real potential to enhance your understanding of the city, create new services, generate jobs and stimulate economic growth. Therefore helping your employees, citizens and businesses take advantage of your data will help kick-start innovation to unlock the potential value from public sector information. This chapter explores the lessons learned from both the **Citadel** pilot and Citadel associate cities to provide recommendations for how to get your data noticed and used.

Governance

Governance of Open Data programs by Cities involves more than just publishing data. There is also a responsibility to ensure it can be accessed and used effectively. **Citadel** aims to help Cities meet this responsibility by adopting an open knowledge approach, providing a collection of data, templates, tools and guides in the form of an Open Data Commons (ODC). The ODC, using the **Citadel** Hub as its focal point, aims to provide opportunities for anyone, regardless of technical background, to be involved in extracting value and knowledge from Open Data. It is regularly updated with FAQs, new tools, and conversations with users as well as links to development sites like GitHub over the course of the project.

The **Citadel** team plans that the ODC will continue to evolve agilely over time based on contributions and good practice from users to become a shared space in the public domain in which cities are use the resources for their own localized ecosystems as well as add back to and enhance them.

When building a local Open Data ecosystem **Citadel** recommends that your bear in mind the following simple but important recommendations from our experience.

If you wait to make sure that everything is error-free both with your data and entire ecosystem then you will never get round to publishing or linking to anything. Many Citadel Associates stated that the reason they had not started an Open Data programme was because they were worried the results would make them look bad. Citadel showcased successful use case stories of the way in which cities partnered with communities to correct inaccurate data sets and enhancing tools to help our Associates overcome fears.

Lesson Learned: Keep the system as open and flexible as possible
Keep seeking and linking to new tools, ideas, people and data in order to keep the ecosystem interesting to provide a continuous learning and experimental platform for innovation. Closed systems close innovation opportunities. Rigorous rules and requirements deter usage.

Lesson Learned: Don't underestimate the need for marketing and promotion
Many Open Data advocates believe that if you simply 'open data people will use it.' The experience of Citadel pilot cities, however, showed that opening data alone is not enough. To generate lasting

value, promotional activities, whether hackathons and workshops or social media campaigns, video and email outreach are needed in order to let people know that data is available and give them the confidence to try to use it.

Empower Your Community

The **Citadel** project started on the premise that the Open Data movement had been dominated by technical actors for too long, leaving many citizens and small businesses left behind. As a result **Citadel** was one of the first projects to create tools to help non-technical people get the same kind of advantages as techies. In the process of creating these tools, Citadel learned the critical value of understanding and addressing the differing needs of key stakeholders.

Lesson Learned: Know your audience's abilities and needs
The wide variety of resources in the first version of Citadel's Open Data Commons (Hub) resulted in a complex online resource that tried to cater for everyone but ended up offering value for no one. The second version of the Hub was then structured according to levels of ability and what people wanted to do with data. Users were now guided to the right tools for their needs. Cities were steered towards data publishing tools, citizens towards easy application generator tools and developers were guided to professional developer environment. A steep upward curve in the use of Citadel was the result.

Understanding your varying audiences and the interplays between them helps to create sustainable processes that will ultimate enable your ecosystem itself to be self-sustaining:

Cities:
- Gain internal buy-in by organising internal workshops to demonstrate the business case for opening data to city leaders and managers
- Ask city managers to set challenges and competitions within the ecosystem to help them engage citizens and developers to use data to solve problems and issues
- Publicize success stories using internal news channels

Citizens:
- Run workshops and hackathon events (business sponsored where possible) to engage the public in the use of your data
- Volunteer their own data collected from their activities
- Provide demos where citizens quickly see examples of data use and be inspired
- Ensure easy access to intuitive support and training tools, e.g. short videos
- Give visibility to created apps on your data portal
- Publicize success stories in local media and promote via social media channels
- Link citizens with schools, universities and businesses within your ecosystem to increase use

Developers:
- Understand professional developers are more interested in the open datasets and 'middleware' tools like convertors than in end-user tools.
- Where possible, be sure to make more and diverse datasets available at regular intervals for this target
- Volunteer their business data to an open portal
- Set professional challenges and competitions to stimulate new data drive business ideas
- Where possible connect the local developer and business communities
- Link your data portal to a Github account to better connect with the developer community and stimulate enhancements and improvements of apps, and tools

Businesses:

- Ask Businesses about what kind of datasets they can contribute to your Open Data portal, and how they can be mashed up with city datasets and those provided by other businesses, and citizens. An example could be a map of spaces and events sponsored by local businesses, together with city attractions and re-seller locations
- Encourage businesses to promote your city's innovation potential by engaging with the development community, sponsoring hackathons, etc.
- Showcase the new innovations and services the businesses unleash with your data.

Lesson Learned: Make your community members feel valued when they make contributions.
You can easily motivate citizens when they generate a new app, colleagues when the open a new data set or developers when the resolve issues on Github by tweeting and blogging about it to give recognition and thanks.

Lesson Learned: Focus on creating a sustainable community where members support each other.
Cities and citizens with little technical knowledge were happy with the outputs of the basic tools on the Citadel website like the App Generator Tool (AGT).. More advanced users, however, found the AGT too basic. To help meet the needs of all skill sets, Citadel established a process wherein developers could use a more advance set of tools on Github to not only generate their own innovations but to also add functionality for more intermediate users.

5. Realising Data's Value: Creating Apps and Services

We need to connect Open Data with those who have the best questions and the biggest needs, a healthcare worker in Zambia, the London commuter traveling home, and go beyond the data geeks and the tech savvy – Dr. Rufus Pollock, Founder, Open Knowledge

The previous sections of this document outlined basic steps any local government needs to take in order to produce high-quality, useful Open Data. However, our experience working with Cities has demonstrated that, as with any process, the Open Data journey has pitfalls and challenges that are unforeseen. To this end, the following section outlines some of the key lessons **Citadel** has learned to stimulate creative people or 'Open Data Makers' to use your data as the foundation of their apps and services.

Overcoming Common Data Usability Challenges

The first set of challenges **Citadel** associates encountered when creating developer-friendly Open Data were around the 'usability' of the datasets they released. Working with more than 500 datasets from 120 local authorities, we were able to identify some common problems that arise when turning Open Datasets into new apps and services:

1. **Data Formats – Citadel** has encountered more than 70 data formats in our work with Local Authorities ranging from widely-used ones like CSV and XML to highly specialist ones used only for specific file types like ShapeFile or Turtle. One of the most important things for a developer using Open data is the ability to get their hands on data in a useful format. While many cities may use XML or XLS for releasing their data, most developers prefer data in JSON or CSV. Creating datasets in the right format to attract developers is a big driver for getting more apps from your datasets.

   To overcome this challenge, Citadel recommends the following, converting your files to developer-friendly JSON files: Visit the Citadel Convertor (http://www.citadelonthemove.eu/en-us/opendata/convertmydataset.aspx) and upload any CSV or Excel file. Within minutes your staff can produce a downloadable JSON file perfect for developers.

2. **Character Encoding** - **Citadel** tools support text that uses the UTF-8 character encoding standard – which has existed since 1996 and has been a world standard since 2003. However, a persistent challenge for our associates was that some popular spreadsheet editing software, including Microsoft Excel and database programmes such as Microsoft Access, still uses regional character encodings (ANSI) by default. These regional encoding issues can lead to special characters such as accents being displayed incorrectly in apps and services.

   To overcome this challenge, Citadel recommends the following fix: Save the excel file as CSV. Now, open the CSV file in notepad and select 'Save As'. At the bottom of the save window you

will see the encoding format set to ANSI. Change the encoding to UTF-8 and save the file[8]. The file will now have universally-encoded characters that can be used by developers.

3. **Geographic Coordinates – Citadel** understands that many of the Open Datasets a city publishes will have location information about the data points. Whether you are releasing data about hospitals, schools, polling stations or public parks, developers need a reliable reference point to display your data on a map. Developers like location data displayed in the form of a Latitude and Longitude value (e.g. 50.856797 4.354781). The reason for this preference is that, unlike address data, Lat/Long are the same everywhere in the world and because all GPS and web mapping software accepts this format. However, from our community of more than 120 Local Authorities, only 15% publish Lat/Long values in their data. Leaving out this information makes a dataset significantly less attractive for developers.
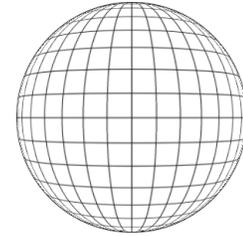
To overcome this challenge, Citadel recommends the following fix: There are many freely available tools that convert address information into Lat/Long. These tools, called 'Batch Geocoders' convert a whole series of values into Lat/Long quickly. Citadel has found the 'Doogal' Geocoder (http://www.doogal.co.uk/BatchGeocoding.php) to be a very simple and effective tool. We recommend copy and pasting your address information into the box and then copy and pasting the resulting list of Lat/Long values back into the dataset. The presence of this information will dramatically increase attraction of your data for developers..

4. **Licenses –** The wrong license on a dataset can be highly detrimental to data re-use by developers - and may lead to the city being criticised for not being truly "Open". Many cities assume that licenses that prevent commercial re-use of information, such as cc-by-nc , help to protect their data from abuse by private companies. However, from our work with companies and developers who specialize in the use of Open Data, we know that such licenses deter developers from using a dataset to build new apps and services, and so greatly diminish the potential benefits to the city.

To overcome this challenge, Citadel recommends that a local authority choose a developer-friendly license like cc-0 (no restrictions) or cc-by (attribution). Open Data under these licenses will be significantly more attractive to commercial developers[9] and will increase the amount of apps and services built using your information.

Overcoming Common Knowledge and Support Challenges

---

[8] For more on this issue please consult the following thread from the Microsoft Help Site: http://bit.ly/1lYCtbY
[9] http://theodi.org/guides/publishers-guide-open-data-licensing

The second set of challenges **Citadel** associates encountered when creating developer-friendly Open Data were around the knowledge and support staff needed to work effectively with Open data. Working with personnel from local authorities across all 6 continents has give **Citadel** a unique insight into some of the common challenges faced by staff working with Open data and some ways to overcome them:

1. **Understanding What Information is Needed –** A common problem among the staff of our associates was that those entering data do not have a full understanding of what information they need to include. Without the right information, a developer will not be able to produce apps and services from the dataset.

To overcome this challenge, Citadel recommends that a local authority create a short memorandum circulated to all staff responsible for the collection and input of data giving them a clear picture of the information that needs to be included. Some of our associates also found it helpful to provide data collection staff with templates that guide them in the type of information they need to provide. In the case of any Point of Interest Dataset, Citadel has produced such a template to allow staff to gather all appropriate information (http://www.citadelonthemove.eu/en-us/opendata/createadataset.aspx)

2. **Citizen-Generated Data -** The explosion of geo-location-enabled smart phones is unleashing a wealth of new citizen-generated data sets which, in turn, is creating a wealth of new opportunities for innovative mash-ups between Open Government Data and other repositories. This development raises a number of questions, however, regarding data quality and maintenance. Work with our Associate network shows that most cities prefer to take a risk averse approach to data that does not come from within their own departments – either by refusing to publish it or requiring onerous maintenance commitments that are likely to deter citizen volunteers and dampen innovation.

In the spirit of open innovation, Citadel believes that an accuracy policy is best promoted by the community itself and by systems that highlight the self-benefit of quality and reliability rather than some form of "enforcement." To help cities get the best of both open government data and citizen volunteered data, we therefore recommend the following:

 I. Use a label system to clearly differentiate between government and citizen-sourced data sets on your publishing portal
 II. Encourage data users and data donors to work together by introducing a rating system or communication tools that facilitate working on data quality in the same way open source communities improve software quality (communally).
III. Aim to develop a certification program that labels apps that use 'official' open data maintained by the city, doing so will not only help to differentiate between government and citizen generated data it will also provide an incentive for app makers to use government data and for civil servants to publish/maintain more data.

IV. Remember that the process of improving the quality and update of both government and citizen generated data eventually leads to APIs at some point to offer data coming from a database, because at certain scale this is where the data is maintained

3. **Updating Procedures –**Developers prefer data sets that are up-to-data and well-maintained as these tend to provide the best basis for production grade app development. Citadel understands, however, that most cities have limited time and resources to dedicate to maintaining all of their data, all of the time. Rather than let this limitation become a barrier to opening data and creating a vibrant ecosystem (even unmanaged data has a higher value than no data), Citadel recommends that cities focus instead on clearly outlining their data maintenance procedures/schedules and labeling data sets accordingly.

At present, only 12% of our 120+ associates have a publicly-accessible policy on updates and maintenance. Yet, this simple measure alone can engender trust in the developer community and boost the attractiveness of your Open Data. To help deliver data maintenance, Citadel recommends 3 steps:

I. Wherever possible, a city should implement an **update schedule for their Open Data** that fits their time and resource constraints. Each dataset will have different update requirements. Some datasets may only require a check every 6 months while others would benefit from more regular attention The important thing is to work smartly. Often staff time is already dedicated to updating this data offline. With the right procedures, maintaining the Open Data can be successfully folded into existing routines at no extra cost.

II. **Use your community** as a maintenance resource. Several of our Associates benefit from engaged groups of citizens who help them maintain the accuracy of their data. To take an example from the UK, the national database of Bus Stop locations (over 300,000 entries) was initially filled with errors when released. The OpenStreetMap community, caught and corrected more than 50,000 errors within 6 months and continue to maintain the accuracy of the dataset free of charge. By engaging your community to help, cities can benefit from an active maintenance policy without the costs or commitments of in-house staff.

III. **Mark Datasets** appropriately according to their maintenance schedule. Developers like to understand where data came from (government, business, citizens) and how often it is maintained. A simple tag by each dataset will let your community know the right information.

4. **Creating an 'Open Data' Culture –** For many of our Local Authorities, Open data remains a confusing and daunting proposition. 41% of our associates had either little or no previous experience of Open Data. Our experience has been that the Open data agenda within a local authority is often driven by one or two 'visionaries' who have been able to successfully open some data, but lack the institutional support to create a systematic programme. Councils that do have the support of their leadership for Open Data can become beacons for data-driven businesses who know they

can rely on a consistent supply of datasets to fuel their work. For those that do not enjoy this support, the failure to create an Open Data culture represents a lost opportunity.

To overcome this challenge, Citadel recommends that a local authority secure the buy-in of senior leadership to the Open Data agenda. Through our work with local authorities, we have learned that the best way to accomplish this is to gather senior leaders together for a day-long workshop on the benefits and potential of Open Data. Once you have secured the backing of these influencers, an Open Data culture can be rolled-out throughout the organization. Citadel has developed an 'Apps4Dummies' workshop concept designed specifically to convince senior leadership of the value of Open Data[10].

---

[10]http://www.citadelonthemove.eu/News/tabid/207/articleType/ArticleView/articleId/82/language/en-US/Open-Data-Builds-Smarter-Cities-New-Open-Data-Toolkit-Fires-Imagination-of-Government-at-London-Workshop.aspx

5. Concluding Summary

Steps for building your Data City include:

1. Before starting your Open Data City strategy understand your business case for opening data from increasing transparency to generating cost savings and unlocking innovation. This knowledge will help you form your own objectives.

2. Once your strategy is in place you should start opening data as soon as possible. Don't wait to achieve perfection. There is no such thing and you and your city will be left behind the technology and policy curve.

3. Follow Citadel recommendations for how, what and where to open data to ensure your data is broadly compliant with industry standards and will be re-usable. The datasets that you open first will be dependent upon the objectives in your strategy.

4. To ensure you unlock the value hidden within your data create a local data ecosystem of tools and guidelines that will help your community use your data. Citadel on the Move contains resources for both novice and experienced developers.

5. Understand your community audience and know their support needs to ensure you engage them and incentivize their participation accordingly - from simple public thank you's and recognition via social media to cash prizes for developing services to solve specific challenges in open hackathon.

6. Be prepared to come up against usability challenges, and utilize your community to help you solve the issues. Citadel has already helped to mitigate common problems:

   a. Use Citadel Convertor to easily create data files that developers can access and use to create new services

   b. Avoid character coding challenges in non-English data sets by saving Excel or CSV files as UTF-8

   c. Use freely available Geo batch coders to add latitude and longitude coordinates to address data so it can be represented in map form using Citadels Application Generator or other visualization tools[11]

   d. Use recognised open licenses so your community is in no doubt that it can use your data for innovation without restriction

   e. Add new data sets to your ecosystem by gaining buy-in from throughout your authority. Use tools like Citadels 'Apps4Dummies' workshop to quickly demonstrate Open Data value to leadership

   f. Share any new developments with your community and the wider Citadel community

7. Remember as a Citadel Associate you are free to draw upon the tools, resources and expertise of the community to help address any problems or questions that you may have. Don't be afraid to talk to us today at www.citadelonthemove.eu

---

[11] Be sure to check the terms and conditions of your chosen geocoder to ensure that your city does not have any conflicts such as 'personal use only' conditions.