DEPARTMENT OF ADMINISTRATIVE SERVICES
BUREAU OF ENTERPRISE SYSTEMS AND TECHNOLOGY
55 Farmington Avenue, Hartford, CT 06105

Connecticut's Open Data Initiative

Category	Open Government & Data, Information and Knowledge Management
State	Connecticut
Contact	Mark Raymond, Chief Information Officer mark.raymond@ct.gov
Project Initiation	February 2014
End Date	December 2015

EXECUTIVE SUMMARY

Connecticut has a culture and history of using data for planning programs and services throughout the State. However, many of these initiatives have been performed by non-governmental entities. As these groups matured and became more sophisticated, they discovered that the quality of available data at the State level was a challenge. In Connecticut, the demand for data was different than larger States. Data users in Connecticut were less interested in developing apps or creating new products, but more interested in policy relevant analysis. There was no shortage of data in Connecticut, and no shortage of users; and there had been numerous attempts to develop data repositories; but those had failed because they all relied on requesting data from State Agencies. In 2014, it became clear that Connecticut needed to change, from the inside – out.

Connecticut's Open Data Initiative was launched in February 2014 through an Executive Order (No. 39) of Governor Dannel P. Malloy. The Executive Order created the position of the Chief Data Officer within the Office of Policy and Management and required the creation of an Open Data Portal (https://data.ct.gov). In addition, the order initiated a 90-day sprint style approach to increasing the amount of data Agencies made available. This included identifying an Agency Data Officer within the first 30-days and identifying and publishing initial datasets within the first 90-days. On March 31, 2014, the Connecticut Open Data Portal went live and contained 100 individual datasets from 20 different State Agencies.

From March 31, 2014 through the end of 2015 Connecticut's Open Data Initiative continued to evolve. The site hosts currently 400 datasets, from all agencies covered by the Executive Order, as well as Connecticut's 4 largest cities, Constitutional Officers including the Secretary of the State and State Comptroller; and the Legislative branch. While other States have created Open Data Portals, Connecticut's initiative was less about the technology and more related to our approach. An Open Data portal itself is only a small portion of the overall implementation of a successful Open Data initiative. Connecticut took a new approach to Open Data that focused on internal structural change. Connecticut's approach focused on flexible governance and value to Agencies, rather than simply shoehorning data into an Open Data Portal. Connecticut's Open Data Hub, an enterprise scale platform that automates the publication of Open Data, ensures the timeliness of our data, and significantly reduces resources required to publish data. Through enhanced outreach to those not required to provide Open Data, Connecticut's efforts made it easier than ever for citizens to access a wide array of data. In essence, Connecticut's efforts around Open Data have created significant, lasting change with regard to the way State Agencies use and value data.

CONCEPT:

The State of Connecticut is home to a diverse and highly skilled and educated population. Connecticut is also fairly unique in that it has no form of County Government, thus many governmental functions and programs that might be performed by Counties are done by State Agencies. As such, State Agencies are subject to frequent and numerous requests for data. While many agencies provide data and information online through their websites, it is often difficult for the public to discover and when they were able to find data it was often locked up in difficult to use formats. This decentralized approach necessarily required an in depth knowledge of state government and its operations. Interested individuals are required to know which agency performs which function or maintains data in order to know where to discover it. Further, there were numerous organizations, including legislators, who were constantly requesting data from State Agencies.

Beyond the public requests for data, was a clear need for State Agencies to share data with one another. Connecticut has numerous interagency task forces, working groups, and committees that all rely upon public data provided by State Government. The groups were often going agency to agency requesting the same data. For instance, data on chronic absenteeism in schools was necessary for juvenile justice issues, educational issues, children's behavioral health planning, and long term economic competiveness. In short, the process of requesting and providing data was redundant and inefficient.

There was no shortage of parties interested in establishing a centralized repository of the State's public data. Prior to 2014, there were several proposals made to the Administration that would establish some form a data portal. Eventually, a non-profit group even began to establish a data portal comprised of State Agency data. This group quickly began to become frustrated with their inability to gain access to data from State Agencies. In essence, Connecticut's struggle to provide greater access to public data had little to do with technology and funding. Connecticut's main problem was lack of leadership, lack of understanding of the utility of agency data, and a lack of structure around the use of data.

With this understanding, Governor Malloy issued Executive Order No. 39 which laid the foundation for Connecticut's Open Data Initiative. The Order:

- Established the position of the Chief Data Officer (CDO) within the Office of Policy and Management;
- Required each agency to designate an Agency Data Officer (ADO). This position is not simply a liaison to the CDO, but rather an upper level manager with broad knowledge of agency operations and data holdings, along with an understanding of the legal and policy issues surrounding the agency's data;
- Required the creation of and Open Data Portal
- Created an open Data Advisory Panel that was composed almost entirely of non-State agency representatives;

- Required agencies to identify easily publishable data within the first 90-days
- Required the CDO to develop and Open Data Policy which would further implement Connecticut's Open Data Initiative.

It quickly became apparent that implementing Open Data in a small State like Connecticut would be different. The majority of the demand for data was not coming from civic technologists looking to develop apps, it was actually a much broader constituency here. In addition, Connecticut was undergoing a significant modernization effort affecting many of our IT systems, which meant that Agencies would struggle to obtain data from legacy systems. Finally, there was an expectation that the State possessed certain critical datasets, which were actually collected locally. Therefore, Connecticut's Open Data strategy focused on the following principles:

- Make it as easy as possible for Agencies to publish data:
  The more effort required of agency personnel, the less likely they are to publish and share data. Agencies needed a variety of options that will align well with their current business practices
- Make it as easy as possible to discover Connecticut's Open Data:
  Most people may not know, nor may they care which agency, branch, or level of government collects which data. They likely have an expectation that key data would be available through a State's Open Data Site, regardless of who collects it. In addition, data should be discoverable not just in the State's Open Data Portal, but elsewhere, including the State's top level website, via search engines, and other data repositories.
- Provide value to State Agencies:
  Agencies need data from other agencies to conduct their business. This requires focusing on their needs as well as the public's. In addition, Agencies want to share their data in meaningful and understandable ways, beyond simply providing bulk raw data. To accomplish this, we need to provide them tools to add context and dimension to their data in a way that makes them feel comfortable; along with the raw data.

With these principles in mind, the Chief Data Officer implemented a broad outreach plan. This plan consisted on regular meetings with Agency Data Officers, meeting with other Elected Officers not covered by the Executive Order (such as the State Comptroller), Legislators, Connecticut's largest cities, the Open Data Advisory Panel, and any other interested parties or Stakeholders.

SIGNIFICANCE:

Connecticut is not the first State to implement an Open Data Portal, or have some degree of an Open Data Policy in place. In many ways, Connecticut was able to learn from other States efforts and develop a solution that would work for our unique structure. However, simply having a mandate and a data portal do not truly comprise an Open Data Program.

Flexible Governance:

Connecticut, like most other State's, implemented Socrata's Open Data Portal solution. We allowed each agency to best determine their approach to publishing Open Data. Some agencies preferred to have multiple data publishers with expertise over a particular dataset or business function. Other agencies (often smaller) preferred to submit data to the CDO for publication in the Open Data Portal. Agencies were allowed to determine the governance model that worked best for their agency, subject to approval of the CDO. In some instances, Agencies were already publishing Open Data. In these cases, agencies simply needed to ensure that location where this data was published was discoverable in the State's Open Data Portal. For example, the Department of Transportation provides a real-time XML feed of highway traffic incidents. Rather than require them to reconfigure an existing business process, the URL and descriptive metadata for that feed was simply made discoverable through the Open Data Portal.

Enterprise Scale Automation:

Long-term sustainability is a key to a successful Open Data Program. After the initial mandate and publication of data, many Open Data initiatives plateau. Datasets grow stale as other mandates and priorities consume personnel resources. Thus, Connecticut needed to ensure that it had an approach to automation. This not only reduces the burden on Agency staff of updating data, but also reduces the effort necessary to extract data from systems. As a result, Connecticut's Bureau of Enterprise Systems and Technology (BEST) developed the Open Data Hub. The Open Data Hub leverages the PilotFish eiPlatform, an enterprise level ETL and data integration platform which the State already owned and could deploy at no-cost. A custom interface was developed that could automate the publication and updating of data from virtually any database platform to the Socrata platform. Several agencies were already using this platform, however BEST implemented a technical assistance model where agencies were eventually able to develop their own interfaces. Further, Agencies are able to utilize a centralized enterprise PiltotFish environment located at BEST, or deploy their own within their agency.

Data Discovery:

Citizens have an expectation that an Open Data site will contain the information they are looking for, regardless of who collects it. Connecticut ensured early on that our Open Data was discoverable and federated with the Federal Open Data site (data.gov). In addition, Connecticut federated with the City of Hartford to ensure that its Open Data was discoverable via the State

repository. Connecticut went on to work with more of our large Cities. Microsites were established for New Haven, Stamford, and Bridgeport; allowing them to publish Open Data at no-cost. This ensures that residents of Connecticut's largest cities have a one-stop-shop for Open Data. Finally, our outreach also went beyond those State Agencies that were not bound by the Executive Order. Data was added from additional Constitutional Offices including the State Comptroller and Secretary of the State; as well as data from the Connecticut General Assembly (legislative branch). Finally, enhancing data discovery is not simply limited to ensuring that data is posted to and discoverable via the Open Data Portal. Connecticut ensured that the search appliance on the state's web portal was configured to return results from the Open Data site. In addition, metadata for datasets in the portal was fine-tuned to ensure discovery by ordinary search engines.

Value to Agencies:

Open Data has value internal to government operation as well as the public. Connecticut established cross-agency working groups, based around specific subject areas such as Health and Human Services, Economic Development, and Emergency Management; in an effort to get agencies around the table to discuss their common data needs. This led to an understanding that there were core datasets that each agency held, that were needed by multiple other agencies (such as locations of schools). Further, the Open Data Portal provided a no cost solution to various agency reporting requirements. For instance, the Department of Economic and Community Develop was required to develop and online tool to provide detailed information on various forms of financial assistance to businesses in the state. They were able to leverage the Open Data Portal to meet this mandate, at no cost. Another example is that the Department of Mental Health and Addition Services was developing a database of pharmacists that prescribe naloxone. The Open Data Portal was leveraged as the mechanism to both collect and display this data. Finally, the State Comptroller was able to sunset their online financial transparency site saving them both money, and staff resources devoted to its maintenance. In short, Connecticut has developed an Open Data ecosystem that benefits not just the public consumers of data, but our agencies themselves.

Beyond this, Connecticut was an early adopter and partner in the creation of Socrata's DataLens feature. This enhanced dashboard style visualization of data allowed our agencies to present data in new ways in which they were already seeking to do, but rarely had the expertise to. Agencies, in turn were also able derive new insight from their own data. For instance, the Chief Medical Examiner's office provide 3 years' worth of overdose related death data. We immediately discovered new trends in both the demographics of these individual, but also the geographic distribution.

IMPACT:

The structure put in place by Connecticut's Open Data initiative has changed the way agencies operate with respect to data. This is embodied in the legislation adopted in 2015 that requires all agencies to enter into data sharing agreements with the State's Office of Policy and Management. Connecticut is now going beyond Open Data, to leveraging all data as strategic asset. State Agencies and the Governor recognize the importance of advancing our use of data to tackle pressing policy issues. All Agencies covered by the Executive Order are publishing data, as well as voluntary participation from the State Comptroller, Secretary of the State, and the Connecticut General Assembly (legislative branch).

Further demonstrating that Connecticut's approach to Open Data has been successful is our top position in the U.S. Open Data Institute's State Data Census. This assessment evaluated states on the availability and openness on 9 critical datasets (while this assessment was not completed until 2016, CT was assessed in November 2015 and retained its top ranking). All of this work is being done without hiring additional staff, and only a single individual, the State's Chief Data Officer, overseeing the work.

Additional Benefits:

- A significant amount of referral traffic comes from news websites and social media sites including Facebook, Reddit, and Twitter; demonstrating the broad awareness and utility of Connecticut Open Data.
- The Open Data site has received approximately 2 million page views and 1.5 million API requests through the end of 2015.
- Of the 400 Open Data sets, 50 are updated on a nightly basis using the Open Data Hub.
- 5 public facing applications have been developed by State agencies using the Open Data site as a data store. Examples include a financial transparency site, and a professional license verification site.
- New user developed products include:
    - The "Your School" website developed by the CT Mirror (an online non-profit news site) that allows people to investigate and compare a variety of data on Schools in Connecticut, beyond just test scores and academic performance.
    - Interactive "Town Profiles" developed by the Connecticut Economic Resource Center and CT Data Collaborative which provide a variety of economic and demographic indicators of CT's 169 municipalities.
- Connecticut was recognized by Good Jobs First as a model in transparency related to data on economic develop assistance programs by providing data on grants, loans, and tax credits to businesses.