

**2009 RECOGNITION AWARDS FOR OUTSTANDING ACHIEVEMENT IN
THE FIELD OF INFORMATION TECHNOLOGY**

Title: Citywide Data Warehouse

Category: Data, Information and Knowledge Management

State Nominating: Washington, DC

B. Executive Summary

Washington, DC's Office of the Chief Technology Officer (OCTO) is spearheading a movement to democratize government data, thereby shifting power from government officials and the halls of bureaucracy to its rightful place--the hands of the citizen. This idea has its roots in ancient Athenian society and draws from the great example of democracy established by its leaders and exercised by its people. In the Agora, or Public Square of ancient Athens, citizens gathered to conduct business, debate civic issues, and drive the decisions of government. Unfortunately, far too often today's citizens are faced with cold, distant bureaucracies, and inaccessible government officials. In the nation's capital, data has been democratized with access to government information made available at any time. OCTO's Citywide Data Warehouse (CityDW), supports the Digital Public Square by supplying live data feeds and access to an organized Data Catalog, which allows the citizen to once again have access to government, hold government accountable, and to make informed decisions that directly shape their quality of life.

Information is believed to be one of the most under-utilized assets in government. Without the proper data sharing in place, government agencies often bear all costs associated with collecting, storing, and protecting data without realizing any noticeable benefits. Like many municipalities, the government of Washington, DC was never known for its transparency of operations. The lack of information made available to the public, and the overall accountability issues created specifically by a lack of publicly available performance data, was exacerbated in the District by a decade of corruption, fiscal crisis, and general mismanagement. The problem was worsened by the sheer volume of permit and service-related requests submitted annually by residents and business owners via phone and e-mail with 413,233 calls and 62,588 e-mails respectively. These requests translated into the need to process over 200 million documents over the years. The increasing need to make this data available to the public became clear. Constituents needed real-time access to information and the transparency this access would afford.

When OCTO became the first and only government agency to catalog and publish raw and near real-time operational data in a variety of formats and in such a large volume, an impact was made on government's ability to engage the District's 588,292 residents. By October 2007, OCTO began publishing more than 200 data feeds through its live catalog and data feeds website, <http://data.octo.dc.gov>. With this publicly shared data, gone were the days when residents and business owners relied solely on e-mail and phone communications, or office visits to navigate a sea of multiple-agency services and information. Not only were constituent services, government transparency, and accountability being provided, but time and money were also being saved. District government employees now spend less time responding to Freedom of Information ACT (FOIA) and other related requests.

C. Description of the Business Problem and Solution

At the current stage of program development, the most significant shortcoming is an absence of data governance and data sharing policies. Sharing data with the public requires clear guidelines on how, when, and with whom data can be shared. As the program has grown, a more formal reusable process was required, and standards and policies have been implemented. OCTO created a standard data sharing agreement that is signed by agency directors and regulates the use of data by the city. The standard data sharing agreement includes all agencies and data owners. The document includes information on the source systems, what data is going to be shared, and with whom. The agreement outlines the service level agreement and who is responsible for supporting the data connection on each side.

By introducing this agreement, OCTO is protecting data owners, which encourages them to be less hesitant to participate in data sharing. This creates a more formal and efficient repeatable process. In 2007 CityDW had zero agreements. By 2008 the program had six major data feeds agreements. With the number of program participants growing, the CityDW team plans to continually develop policies and standards to drive both technical guidelines and business process that are accepted and mandated across all 86 District agencies and departments.

OCTO started the CityDW program to design a data warehouse to store agency data, create various presentation tools for managers, line workers and the public, and to increase government transparency by providing more information across agencies and to the public. By October 2007 the new Data Catalog was released to the public, providing more than 150 data feeds in five formats.

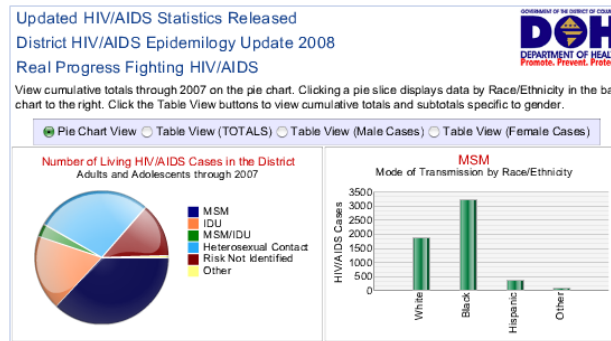
Today applications provided by CityDW include (but not limited to):

- Neighborhood View and Mobile View (designed for mobile devices), analytical tools with strategic agency datasets, for referencing information about crime, service requests, inspections, ownership, property, tax, and other related data. In the past, researching ownership information has been time consuming. These applications have reduced the time and effort spent through the process of investigating properties and their owners. Additionally, Neighborhood View allows employees to view data on maps and drilldown into the details of individual data records.

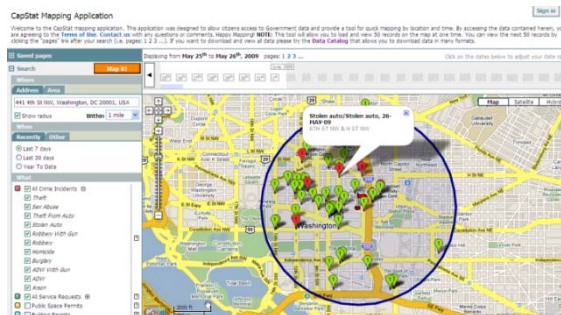


- CapStat Daily Management Report – summary report with up to date financial, procurement and personnel information sent to agency executives every morning

- Interactive dashboards highlighting the most important issues government is currently dealing with.



- Summary reports website providing the prepared reports on multiple topics
- Digital Public Square
- Data Catalog (<http://data.octo.dc.gov/>)-This is the most ambitious transparency program, launched in October 2007 recently duplicated in May 2009 by the federal government through data.gov. Residents and government workers can access raw data in XML, Text/CSV, KML or special geo formats. Crime incidents, service requests, contracts and spending can be accessed by most desktop spreadsheet applications with just a few clicks, or viewed in Google Maps or Google Earth in one click. The full Data Catalog provides over 278 datasets.
- CapStat Mapping application provides residents with the ability to choose the location and see info relevant to their everyday life – crime, permits, service requests and schools.



- Internal ad-hoc reporting environments are used by multiple city agencies on different levels. CityDW insures the consistency of information and reporting – a citizen, an agency line worker, a city executive see the same data.
- Data streams are provided within the government. For example, Calls for service (the Mayor’s 311 Call Center) and crime data is used by security and emergency management agencies to isolate areas that need the most attention. If there are a lot of calls for trees down on a block or in a neighborhood, the Homeland Security and Emergency Management agency can direct DDOT to that location.

D. Significance of the Project

The most significant achievement of DC's Data Feeds is its increase of citizen participation. For example, for an investment of only \$50,000, the District's Apps for Democracy contest challenged residents to utilize information available from DC's Data Feeds to develop consumer-based applications. Local developers produced 47 applications that were conceived, developed, and delivered in 30 days (www.appsfordemocracy.org). Winning applications include a DC biking guide, government purchases over \$2,500, parking meter locations, community garden sites, and more. By treating citizens as co-creators of government, the District secured software worth millions of dollars that the government alone would not have been able to create, particularly in this economic climate. The CTO estimates it would have cost \$2.6 million dollars to develop the applications in-house. The District has already put seven of the applications in use since the contest concluded on November 12th. In fact, a few of the 47 applications such as DC Historic Tours and ILiveAt are highlighted on the official DC government inaugural website; providing 588, 292 residents and a projected 2 million visitors with creative tools that provide customized information on inauguration activities and DC attractions.

Another example is a talented District resident who provides daily updates about her neighborhood on www.Jdland.com. This resident has created an intricate website that uses the District's data feeds to provide real-time economic development information, including permit status and crime statistics. This webmaster's audience of neighbors, residents, DC government officials, commercial real estate professionals, planning and development professionals, and baseball fans appreciate the depth of information catered to their interests. As a result, jdland.com won a Knight-Batten Citizen Media for Innovation in Journalism Award in September 2008.

As a result of making data feeds available for public use, the District has also seen a significant decrease in Freedom of Information Act (FOIA) requests. From the time the District deployed data feeds in 2006, the total number of FOIA requests decreased from 15, 527 in FY2006 to 8, 280 in FY2007. A *Washington Post* reporter explained, "DC government also benefits" from providing open access to this information. Expounding on this, this reporter further noted that, "A lot of times we can answer our own questions and realize that something is or is not worthy of a story....Having datasets available for download rather than requiring the much more laborious Freedom of Information Act (FOIA) process also saves the government a lot of time."

The beauty of the District's Data Catalog and Feeds program is that it is many things to many people. As cited in these examples, the Data Catalog is utilized by and benefits both government and the public, and has limitless applications and uses that are far-reaching and ultimately serve the needs of all.

E. Benefit of the Project

Over time, CityDW's implementation strategy has evolved significantly based on the requests of its customers. The vast degree of transparency afforded by the Data Catalog's data feeds allows both citizens and government agencies to question, participate on a scale never seen before, and in the end, request even more data. With this level of involvement, as well as the constant requests for additional data, the CityDW program has learned to shift and adapt accordingly to meet customers' needs while never losing sight of its mission to provide operational data that provides government transparency, accountability, and ultimately places power in the hands of the citizen.

Some of the program's most notable benefits are:

- A. Cost efficiency: The program provides one centralized location for storing data.
- B. Information in the data warehouse is under the control of the data owner; so, multiple copies of truth are not needed.
- C. The program is separate from agency systems; so, there is no slowing down of operational systems.
- D. CityDW created pre-defined common geographies (FIAs, Wards, PSAs) used enterprise-wide.
- E. From CapStat sessions, the Mayor made and has tracked commitments to create 10,000 affordable housing units, institute focused improvement areas to fight crime, and reduce the city's fixed costs. This real-time access to data allows the Mayor to hold agency directors accountable.
- F. Data is made available quickly and CityDW provides frequent updates in some cases as often as every 15 minutes.
- G. In October 2008 The *Apps for Democracy Contest* ran for 30 days, which provided an opportunity for unforeseen cooperation between the public and DC government. By inviting local developers to use government data, a true partnership between government and its citizens proved to be mutually beneficial for both parties. There were 47 applications produced for the cost of \$50k; an estimated cost savings of more than \$7 million.