



Montgomery County Data Innovation Committee

Leslie Rubin

Montgomery County Data Innovation Committee

OLO Report Number 2014-4

December 3, 2013

Over the past decade, the scope and content of data related to government activities has changed dramatically. The sheer quantity of data available for public consumption, the way in which it is structured and how datasets are used has the potential to impact program planning, analysis and evaluation at the local government level.

This report responds to a request from the Council to examine the feasibility of creating an advisory board with the expertise to conduct large scale data analytics.

Data and Data Analytics

Data analytics refers to “a process of inspecting, cleaning, transforming, and modeling data with the goal of discovering useful information, suggesting conclusions, and supporting decision making.” A current buzzword in the world of data analytics is “Big Data” – referring to the vast quantities of data that are created in today’s world by businesses, local governments, and individuals.

Many reports estimate that the amount of data created annually is expected to double every year. Data come from sources as varied as financial transactions, emails, sensor readings, and social media, and often are too big and diverse to analyze with traditional data analysis tools. Business experts and academics contend that analyzed and used strategically, data can help government agencies:

- Identify and reduce inefficiencies,
- Improve productivity,
- Reduce security threats and crime, and
- Create or improve transparency.

In jurisdictions that use advanced analytics, many use the data for predictive analytics – to look forward and predict future events or behavior to help shape decision-making. For example:

New York City combined data on illegal property conversion complaints, foreclosures, tax liens, and neighborhood demographics and found that certain factors correlate with a high risk of residential fire. A team then created a risk assessment model to prioritize illegal conversion inspections. Before creating the model, inspectors found seriously hazardous fire conditions in 15 percent of inspections. Using the model, inspectors found seriously hazardous fire conditions in 75 percent of inspections.

Data Use in the County Government and County Agencies

The County Government and County agencies generate and maintain large amounts of data. And, in the past year, the Executive has launched dataMontgomery – a platform for publicly publishing disaggregated County Government datasets.

A few County Government departments are exploring ways to use complex data analytics or predictive analytics to further their missions. No departments currently use these advance practices regularly.

Montgomery County Public Schools (MCPS), Montgomery College, and the Montgomery County Planning Department all use advanced data analytics, in one form or another, to aid decision-making. MCPS, for example, created a model using attendance, behavior, and course failure data to identify students with an increased likelihood of dropping out of school – to help target resources. Additionally, the Planning Department combined various data to predict where demand for bicycle use will be greatest – to help prioritize where to build bicycle routes and other relevant facilities.

Montgomery County Data Innovation Committee

If the Council would like to pursue creating an advisory board to conduct large scale data analytics, this report describes how OLO would recommend structuring a **Data Innovation Committee** to advance the use of innovative data analytics in the County Government and County agencies.

The key components of the proposed structure for a Data Innovation Committee are:

Data Innovation Committee Members	10-15 data analysis experts
Data Innovation <i>Executive</i> Committee	Separate from the Data Innovation Committee – this group would monitor the work and progress of the Data Innovation Committee, develop an annual work program, and recommend individuals to the County Council to sit on the Data Innovation Committee
<i>Executive</i> Committee Members	Representatives from: Office of Legislative Oversight County Council County Executive Rotating between MCPS, Montgomery College and the Planning Department Data Innovation Committee Chair Data Innovation Committee Vice Chair
Data Innovation Committee Work Program Input	From: OLO, County Council, County Executive, MCPS, Montgomery College, Montgomery County Planning Department
Selection of Data Innovation Committee Chair and Vice Chair	Recommended by the other Executive Committee representatives, approved by the Council
Selection of Data Innovation Committee Members	Recommended by the full Executive Committee, approved by the Council
Terms of Service	2 years, with the possibility of reappointment by the Council
Data Privacy	Ensure compliance with federal, state, and local privacy laws

Other Considerations. If the Council establishes a Data Innovation Committee, OLO strongly recommends taking steps to ensure that the Committee remain free from both internal and external political pressures to avoid hindering agency willingness to participate in the Committee.

Timing of Council Consideration

In December 2012, the Council enacted Bill 23-12 on open data. The legislation requires the Chief Administrative Officer to develop an Open Data Implementation Plan within 18 months that describes public datasets under the control of the County Government and that prioritizes publication of the datasets on the County Government’s open data portal. Executive Branch representatives report that the Open Data Implementation Plan will be completed by the summer of 2014.

In order to have a fuller understanding of available County Government data, OLO recommends that Committee and Council formation of a Data Innovation Committee coincide with discussion of the Open Data Implementation Plan when the Chief Administrative Officer releases it next summer.

Office of Legislative Oversight Report 2014-4

**MONTGOMERY COUNTY
DATA INNOVATION COMMITTEE**

Table of Contents

Executive Summary i

I. Authority, Scope, Organization 1

II. Data 3

III. Data in Montgomery County Agencies 7

IV. Montgomery County Data Innovation Committee 11

V. Examples of Data 15

VI. Agency Comments 18

Appendix

Chapter I. Authority, Scope, and Organization

A. Authority

Council Resolution 17-830, *FY 2014 Work Program for Office of Legislative Oversight* (July 30, 2013).

B. Scope, Purpose, and Methodology

Over the past decade, the scope and content of data related to government activities has changed dramatically. The sheer quantity of data available for public consumption, the way in which it is structured and how datasets are used has the potential to impact program planning, analysis and evaluation at the local government level. A 2013 survey of state and local government officials revealed, however, that only two percent of jurisdictions have a strategy to use advanced data analytics effectively.

Recent examples of jurisdictions using data analysis to improve services include:

- A county used predictive analytics to monitor water consumption, identify leaks, and predict how much water would be needed, where and when. The county reduced water use by 20 percent, saving \$1 million a year.
- A city combined data on illegal property conversion complaints, foreclosures, tax liens, and neighborhood demographics and found that certain factors correlate with a high risk of residential fire. A team then created a risk assessment model to prioritize illegal conversion inspections. Before creating the model, inspectors found seriously hazardous fire conditions in 15 percent of inspections. Using the model, inspectors found seriously hazardous fire conditions in 75 percent of inspections.

This report responds to a request from the Council to examine the feasibility of creating an advisory board with the expertise to conduct large scale data analytics. Accordingly, the report:

- Examines how data is generated today;
- Describes the use of data analytics and predictive analysis by other jurisdictions to improve efficiency and services;
- Examines the County Government's and outside agencies' practices for collecting and analyzing data; and
- Describes how the Council could create a "Data Innovation Committee" to undertake projects involving advanced data analytics and/or predictive analytics.

C. Organization of Report

Chapter II, Data, describes the types and amount of data generated today and how some state and local jurisdictions and school systems are using data to help advance their core missions.

Chapter III, Data in Montgomery County Agencies, summarizes how the County Government, M-NCPPS, Montgomery County Public Schools, and Montgomery College use data.

Chapter IV, Montgomery County Data Innovation Committee, describes a model for a Data Innovation Committee consisting of experts in data analytics to increase the use of innovative data analytics in the County Government and County agencies.

Chapter V, Examples of Data, briefly summarizes different types of datasets found at the federal, state, and local levels.

Chapter VI, Agency Comments, includes the Executive Branch's and agencies' comments on the final draft of the report.

D. Acknowledgements

OLO received a high level of cooperation from those who contributed to this report. OLO appreciates the information shared and insights provided by all who participated. In particular, OLO would like to acknowledge the time and expertise of the following individuals:

- Benjamin Bederson, Special Advisor to the Provost on Technology and Educational Transformation, University of Maryland
- Larry Bowers, Chief Operating Officer, Montgomery County Public Schools
- Sherwin Collette, Chief Technology Officer, Montgomery County Public Schools
- Christopher Corcoran, Office of Policy and Strategic Planning, City of New York
- Allison Druin, Chief Futurist, University of Maryland
- David Gottesman, Manager, CountyStat, MCG
- Ashley Hand, Chief Innovation Officer, Kansas City, MO
- Katherine Hillenbrand, Research Assistant, Data-Smart City Solutions, Harvard Kennedy School
- Daniel Hoffman, Chief Innovation Officer, MCG
- Fariba Kassiri, Assistant Chief Administrative Officer, MCG
- Dieter Klinger, Chief Operating Officer, Department of Technology Services, MCG
- Angela Lindsay, Crime Analyst Supervisor, Montgomery County Police Department, MCG
- Kevin Long, Senior Planning and Policy Analyst, Montgomery College
- Robert Lynch, Director, Office of Institutional Research & Analysis, Montgomery College
- Susan Madden, Chief Government Relations Officer, Montgomery College
- Christopher McGovern, GIS Manager, Montgomery County Planning Department
- Roberto Ruiz, Research Manager, Montgomery County Planning Department
- Sonny Segal, Chief Information Officer, MCG
- Michael Tarquinio, Enterprise Architect, Department of Technology Services, MCG
- Kathleen Wessman, Vice President, Planning and Institutional Effectiveness, Montgomery College

Chapter II. Data

Data analytics refers to “a process of inspecting, cleaning, transforming, and modeling data with the goal of discovering useful information, suggesting conclusions, and supporting decision making.”¹ A current buzzword in the world of data analytics is “Big Data” – referring to the vast quantities of data that are created in today’s world by businesses, local governments, and individuals.

Many businesses, academics, and data scientists believe that an organization’s ability to “mine” and analyze data, Big Data in particular, is key to future growth and development. Few state and local governments, however, have begun to explore the possibilities associated with advanced data analytics. One survey found that 44 percent of state and local governments are not talking about Big Data analytics and another 39 percent are just beginning to discuss how Big Data could impact their agencies.²

This chapter explains what “Big Data” is and what makes data “Big,” followed by a description of how some state and local governments and school systems are using advanced data analytics to drive their decision-making.

A. What is Big Data?

Big Data refers to “a phenomenon defined by the rapid acceleration in the expanding volume of high velocity, complex, and diverse types of data.”³ In plain(er) English, Big Data refers to the vast amount of raw digital information generated and collected today.⁴ These data come from sources as varied as financial transactions, emails, sensor readings, and social media, and often are too big and/or diverse to capture, store, and/or analyze with traditional data analysis tools.

Data experts describe Big Data in terms of three (or sometimes more⁵) “Vs”:

- **Volume:** refers to exponentially increasing amount of data generated and collected today.
- **Velocity:** refers to the increasing rate at which data is created today.
- **Variety:** refers to the different types of non-traditional data created today.

Volume. The volume of data created around the world today almost defies comprehension. For example, the amount of data created in 2011⁶ would fill 200 billion, 2-hour HD movies, which one person could watch for 47 million years straight.⁷ Similarly, the amount of U.S. healthcare data in 2009⁸ was 30 times larger than the data capacity needed to contain all words ever spoken by human beings on earth.⁹

¹ http://en.wikipedia.org/wiki/Data_analytics.

² *The State and Local Big Data Gap*, MeriTalk, at p. 3 (Apr. 29, 2013) [hereinafter “*Big Data Gap*”].
<http://www.meritalk.com/state-and-local-big-data.php>.

³ *Demystifying Big Data: A Practical Guide to Transforming the Business of Government*, TechAmerica Foundation: Federal Big Data Commission, at p. 7 (Oct. 2012) [hereinafter “*Demystifying Big Data*”].
<http://www.techamerica.org/Docs/fileManager.cfm?f=techamerica-bigdatareport-final.pdf>.

⁴ *Ibid.* at p. 9.

⁵ Among others, data experts also talk about (1) “veracity”: or trustworthy data, (2) “visualization”: referring to the importance of data presentation, and (3) “value”: referring to the relative important of the data to an organization. *Big Data Big Promise*, The Center for Digital Government, at p. 6-7 (2013). <http://www.govtech.com/library/>.

⁶ 1.8 zettabytes of data was created in 2011.

⁷ *Demystifying Big Data* at p. 9.

⁸ There was 150 exabytes of U.S. healthcare data in 2009.

⁹ *Demystifying Big Data* at p. 9.

While the above examples describe data at a global or national level, one study estimates that the average state and local government agency stores 499 terabytes of data – equal to 10 million four-drawer filing cabinets of text.¹⁰

Velocity. Data today are created at lightning speed. People around the globe send 150 billion new email messages every day – illustrating both the tremendous amount of data creation (**volume**) and highlighting the speed at which it is being created (**velocity**).¹¹ Many reports estimate that the amount of data created annually is expected to double every year.¹² Technology today now allows people to capture and analyze data immediately – in real time.¹³

Variety. A significant challenge to harnessing the potential of data stems from the types – or variety – of data collected today that did not exist 10 or 15 years ago. As one report simply explains:

In the past, most data was transactional in nature and structured. It had the same pre-defined formats and lengths that easily fit into fixed fields in relational databases or the columns and rows of spreadsheets. It was fairly simple to compare, organize and analyze.¹⁴

These data are referred to as “structured” data. While structured data (data that fit well into relational tables) such as transactional business data, still make up a chunk of local government data, governments now generate and collect numerous types of “unstructured data,” such as:

- Emails,
- Video footage,
- Photographs,
- Audio recordings,
- Geographic/map data,
- Social media data, and
- Data from sensors (monitoring, e.g., traffic speed, weather, environmental surroundings).

Various reports estimate that 80 to 90 percent of data generated today is unstructured.¹⁵

¹⁰ *Big Data Gap* at p. 5. As a local example, in FY12, Montgomery County Government computer servers received 85 million emails.

¹¹ Reynolds Lewis, Katherine, “What the heck is Hadoop?”, FCW (Mar. 25, 2013). <http://fcw.com/articles/2013/03/25/what-is-hadoop.aspx>.

¹² *Demystifying Big Data* at p. 9.

¹³ *Big Data Big Promise* at p. 6.

¹⁴ *Ibid.* at p. 7.

¹⁵ *Proactive Planning for Big Data*, CDW-G, at p. 2 (2013). <http://www.fedtechmagazine.com/sites/default/files/122210-wp-big-data-df.pdf>. *Big Data, Big Promise*, at p. 7.

B. What Can Data Analytics Do?

The purpose of data is not simply to collect it, but to put it to use to solve problems. One report asserts:

Big data isn't a magic bullet – those are in short supply – but it presents very real opportunities to transform public service by driving dramatic improvements to both citizen-facing services and internal workflow. Federal, state and local governments alike are sitting on a treasure trove of information, but often don't know what to do with it or how to use it.¹⁶

In the summer of 2013, the Harvard Kennedy School established Data-Smart City Solutions, to help jump start the use of data and data analytics in local government. The program highlights successful examples of data use at the local government level, identifies best practices, and supports jurisdictions seeking to begin data projects.

Many business experts and academics contend that analyzed and used strategically, data can help government agencies:

- Identify and reduce inefficiencies,
- Improve productivity,
- Reduce security threats and crime, and
- Create or improve transparency.¹⁷

In jurisdictions that do mine their data, many use those data for predictive analytics – to look forward and predict future events or behavior to help shape decision-making. The examples below describe how federal, state, and local governments and local school systems have put tools in place to effectively analyze and use data.

The Federal Centers for Medicare & Medicaid Services created a Fraud Prevention System that uses predictive modeling technology and screens all Medicare fee-for-service claims to detect possible improper payments. In 2012, the system's first year of implementation, it stopped or flagged \$115.4 million in improper payments, saving an estimated \$3 for every \$1 spent on the system.

New York City's (NYC) Department of Environmental Protection wanted to crack down on restaurants' illegal dumping of cooking oil into sewers, causing over half of all clogged drains in NYC. The City's Office of Policy and Strategic Planning compared City data on certifications of restaurants with carting services to take away their grease (a legal requirement) with geo-spatial data on sewers and identified a list of probable illegal dumpers. Using this list, inspectors had a 95 percent success rate in tracking down the illegal dumpers.¹⁸

Miami-Dade County used predictive analytics to monitor water consumption, identify leaks, and predict the amount of water needed at a given location at a given time. This analysis resulted in a 20 percent decrease in water consumption and an annual savings of \$1 million.¹⁹

¹⁶ *Big Data Big Promise* at p. 2.

¹⁷ *Ibid.* at p. 3; *Demystifying Big Data* at p. 13-15.

¹⁸ Alan Feuer, "The Mayor's Geek Squad," *The New York Times* (March 24, 2013).

http://www.nytimes.com/2013/03/24/nyregion/mayor-bloombergs-geek-squad.html?pagewanted=all&_r=0.

¹⁹ Rachel Wheeler, "How Big Data Can Benefit Local Governments," Experian (2013). <http://www.qas.com/data-quality-news/how-big-data-can-benefit-local-governments-9571.htm>.

The **State of Michigan's Department of Community Health** combined 15 separate databases with health-related data into one data warehouse. The warehouse included data on Medicaid, community mental health, maternal and infant health services, and vital records, among others. Combining these data let the State track clients, spending, and outcomes across programs, and using data analytics, the State was able to improve outcomes in numerous areas:

- The State doubled its rate of Medicaid fraud identification.
- The State jumped from last to first in the United States for child immunizations.
- The State discovered that 80 percent of all lead poisonings occurred in 14 communities and targeted prevention efforts in those communities – ultimately reducing incidence of childhood lead poisoning by 35 percent.
- The State identified Medicaid-enrolled children at high-risk for flu complications so their doctors could give them flu shots.²⁰

The **Charlotte-Mecklenburg County, NC School District** uses data mining to identify and provide special assistance to at-risk students. The district created a predictive model that includes “truancy, disciplinary problems, changes in course performance, and overall grades” and found that they can predict with “reasonable probability” students who are likely to drop out of school.²¹

²⁰ *Big Data Big Promise* at p. 12-13.

²¹ Darrell West, *Big Data for Education: Data Mining, Data Analytics, and Web Dashboards*, Brookings Institution, at p. 5 (Sept. 2012).

<http://www.brookings.edu/~media/research/files/papers/2012/9/04%20education%20technology%20west/04%20education%20technology%20west.pdf>.

Chapter III. Data in Montgomery County Agencies

County Government agencies generate and maintain large amounts of data. This chapter briefly describes the types of data maintained by the County Government, Montgomery County Public Schools, M-NCPPC, and Montgomery College and outlines whether and how the County agencies are using advanced data analytics and predictive analysis.

A. openMontgomery and County Government Data Analysis

County Government departments and offices generate and/or maintain great amounts of data. In the past year, the County Government has taken proactive steps to begin making its data available to the public – with the Council enacting legislation governing the release and publishing of data¹ and the Executive Branch launching openMontgomery, an initiative “for the future of digital government.”² One of the four components of openMontgomery is dataMontgomery – a platform for publishing disaggregated County Government datasets.

As of November 5, 2013, the Executive Branch had published 34 datasets on the dataMontgomery website (<https://data.montgomerycountymd.gov/>). These datasets range from County Government employees’ salaries to restaurant inspection data to demolition permit applications and work status.

Department of Technology Services (DTS) staff are in the process of meeting with County Government department staff to identify all datasets that departments have created and expect to complete this inventory by the summer of 2014. As part of the openMontgomery initiative, DTS also is updating a 2008 catalogue of the IT systems that County Government departments use to conduct their work. DTS staff have identified over 700 systems so far (many of which are utility programs, such as anti-virus programs, that don’t generate data).

Once staff have identified the County Government’s datasets, they will continue publishing them on dataMontgomery. Staff in the Office of the County Executive have created a list of criteria that they will use to determine the order in which to publish datasets. The criteria include factors such as how useful the data is, whether the data contains privileged information that cannot be released to the public, the staff time required to keep datasets up-to-date, and the cost of publishing datasets.

Data Analysis in Montgomery County Government. The Office of the County Executive has been leading efforts to increase the use of data analysis in County Government departments. Representatives report that a few County Government departments have approached the idea of using complex data analytics and/or predictive analysis to support their missions. No departments currently use these advanced practices routinely.

The Montgomery County Police Department’s (MCPD) Crime Analysis Section within the Criminal Investigative Division tracks data on major crimes in the County (e.g., murder, rape, aggravated assault, sex offenses) and has data going back to 1997. The Crime Analysis Section uses data for tactical awareness (to provide officers relevant information for when they are policing) and for investigative support. The section also supports MCPStat, which reviews trend data. MCPD staff report that the amount of crime in the County is too low to make efforts at predictive analysis useful or worthwhile.

DTS representatives report that the department does have some open source tools available to facilitate large-scale data mining.

¹ See Bill 23-12, Administration – Open Data (enacted Dec. 4, 2012).

² openMontgomery: Montgomery County Maryland’s Digital Government Strategy, at p. 1 (Dec. 12, 2012).

B. Data Analysis in County Agencies

The Montgomery County Planning Department, MCPS, and Montgomery College each use data and data analytics to facilitate their missions. In addition, the agencies all use predictive analytics to some extent to guide future decision-making. The sections below describe data use and analysis in these three agencies.

1. Montgomery County Planning Department

The Montgomery County Planning Department manages a vast amount of data and information and makes much of it available to the public. Examples of just some of the data available from the Department include:

- Data on the parcels and zoning of land in the County,
- All County addresses,
- The status of and documents related to development applications,
- County-wide intersection counts, and
- Special Protection Areas that require protections beyond standard environmental laws and regulations for land development and certain uses.

The Department's data serve both as a resource for internal decision-making and as a source of information for residents and others interested in Montgomery County land use. The Department engages in some predictive analysis to help facilitate decision making by County Officials. For example, Department staff calculate the number of students living in an average dwelling (such as a single family home or an apartment). Montgomery County Public Schools subsequently uses these data to project school enrollment based on changes to the housing supply.

In 2011, Department staff generated a bicycle heat map – predicting where demand for bicycle use will be greatest – to help the Planning Board and other decision-makers prioritize where to build bicycle routes and other relevant facilities. Department staff used 11 weighted factors – such as residential and employment density, proximity to mass transit stops (Metro and MARC), and proximity to schools – to create the map.

In addition to the Planning Board and other local government decision-makers, Planning Department staff report that other frequent users of the Department's data include real estate developers, who, for example, can track existing and planned County development to identify areas that may be underserved by commercial or residential developments.

2. Montgomery County Public Schools

Montgomery County Public Schools (MCPS) collects and uses many types of data on a day-to-day basis. These data include:

- Student performance and demographic data,
- Data on student suspensions,
- Data on serious incidents in schools (e.g., fights, fires, equipment malfunctions, computer misuse, bullying, etc.), and
- Data on facility utility usage recycling rates.

In the past several years, MCPS has analyzed data to help reduce operating costs in several areas:

- MCPS monitors and reduces electricity use during peak summer hours through its Peak Load Management initiative to reduce “system capacity charges” that total 10 to 15 percent of total electricity costs. Through the initiative, MCPS saved between \$1.4 million to \$1.6 million a year between 2010 and 2013 – compared to costs in 2007 (the baseline year).
- Through efforts to increase recycling, MCPS saved between \$215,000 and \$245,000 each year between 2009 and 2012 in fees associated with the disposal of solid waste – compared to costs in 2005 (the baseline year).
- Through analysis of water fixtures, general use, irrigation systems, and best management practices, MCPS staff reduced water use in high schools – MCPS’ largest facilities and largest consumers of water per square foot – by over 81 million gallons per year since 2008.

In addition to using data to analyze and reduce operating costs, MCPS’ Office of Shared Accountability released a report in March 2013 that analyzed the relationship between attendance, behavior, and course failures to identify students with an increased likelihood for dropping out of school.³ The report recommends developing a monitoring tool to identify students with an increased likelihood of dropping out in order to target intervention strategies toward these students.

MCPS makes a variety of datasets available on its website. The website for the Office of Shared Accountability⁴ provides reports and data on student performance, program evaluations, enrollment, and other topics.

3. Montgomery College

Montgomery College collects and uses data in a variety of ways. Many State and Federal requirements compel the College to compile and report data – such as requirements for Federal financial aid, State accountability benchmarks, and facilities’ energy use. The College also uses data to facilitate its mission and aid in decision-making. Examples include:

- Monitoring and evaluating the success of academic programs,
- Assessing local workforce needs to identify opportunities for new classes or programs,
- Administering a biennial Community College Survey of Student Engagement used to identify areas for improved student engagement, and
- Analyzing the use rate of classrooms to compare use to the projected enrollment in the College’s Facilities Master Plan.

The Office of the Vice President for Planning and Institutional Effectiveness⁵ oversees the College’s primary data collection and analysis. Its mission includes:

- Coordinating the College Area Review and Outcomes Assessment processes.
- Leading long-range, strategic, and tactical planning processes.
- Facilitating policies that ensure college data integrity.
- Interpreting data to inform and enhance decision making.⁶

³ http://montgomeryschoolsmd.org/departments/sharedaccountability/reports/2013/Just%20the%20Right%20Mix_MCPS_West2013.pdf

⁴ <http://www.montgomeryschoolsmd.org/departments/sharedaccountability/>

⁵ <http://cms.montgomerycollege.edu/edu/departments.aspx?id=4734>

Within the Vice President's Office, the College's Office of Institutional Research & Analysis provides information, data, and analysis to Federal, State, and County agencies, accreditation agencies, Montgomery College administrators, and newspapers and college guidebooks, among others, and publishes some data on its website.⁷ Examples of the Office's publications include reports on enrollment, campus and distance learning, and student progress and success; and surveys and evaluations.

The College also engages in some predictive analysis to facilitate its missions. For example, it compiles data to project both future enrollment and faculty and facilities requirements.

⁶ Ibid.

⁷ <http://www.montgomerycollege.edu/Departments/inplrsh/>

Chapter IV. Montgomery County Data Innovation Committee

Over the past decade, the scope and content of data related to government activities has changed dramatically. The sheer quantity of data available for public consumption, the way in which it is structured, and how these datasets are used has the potential to impact program planning, analysis and evaluation at the local government level.

Many Montgomery County Government departments analyze data – either as a means to further their core mission (e.g., the Department of Corrections) or as part of their core mission (e.g., OLO, CountyStat). At the same time, some state and local jurisdictions have taken steps to expand the use of their data through advanced data analytics.

This report responds to the Council’s request to examine the feasibility of creating an advisory board with the expertise to conduct large scale data analytics. If the Council would like to pursue this strategy, this chapter describes how OLO would recommend structuring such a group - a Data Innovation Committee.

A committee, as described below, may help facilitate an increase in the use of innovative data analytics in the County Government and, if applicable, in the County agencies. Interpreting and understanding results from advanced data analytics can help jurisdictions enhance decision-making, increase efficiency, and/or improve outcomes.

The structure of the Data Innovation Committee described below is based on similar ad hoc groups developed at the federal, state, and local levels – created for the purpose of pursuing innovation in government. While these groups have very different missions from the Committee described in this chapter, the groups illustrate ways that different jurisdictions tap outside expertise in innovative ways. Examples of such groups include:

- The President’s Council of Advisors on Science and Technology,¹
- The Federal CIO Council’s Innovation Committee,²
- The Federal Leading Education by Advancing Digital (LEAD) Commission,³
- The New Hampshire Governor’s Commission on Innovation, Efficiency & Transparency,⁴ and
- The Kansas City, MO Mayor’s Challenge Cabinet.⁵

See Appendix A for a brief description of each group.

Montgomery County Data Innovation Committee

The description below describes one way to establish a Data Innovation Committee made up of a cross-sector group of experts in data analytics to examine and interpret data in new ways on behalf of the County Government (and County agencies).

Committee Structure and Executive Committee. The Data Innovation Committee, itself, could consist of ten to fifteen subject matter experts with an appointed Chair and Vice Chair of the Committee. If the Council would like to pursue this project, OLO also recommends establishing an Executive Committee to monitor the work and progress of the Committee, develop the Committee’s annual work program, and recommend individuals to the County Council to sit on the Data Innovation Committee. The Executive Committee could include:

¹ <http://www.whitehouse.gov/administration/eop/ostp/pcast/about>

² <https://cio.gov/about/groups/innovation-committee/>

³ <http://www.leadcommission.org/about-lead>

⁴ <http://www.governor.nh.gov/media/news/2013/pr-2013-05-29-innovation-commissio.htm>

⁵ <http://kcmayor.org/newsreleases/mayor-james-announces-appointments-to-his-challenge-cabinet>

- A representative from OLO,
- A representative from the County Council,
- A representative or representatives from the Office of the County Executive (e.g., the Chief Innovation Officer and/or the Director of CountyStat),
- A representative rotating from year-to-year between MCPS, Montgomery College, and the Montgomery County Planning Department,
- The Data Innovation Committee Chair, and
- The Data Innovation Committee Vice Chair.

The representatives on the Executive Committee from OLO, the Council, the County Executive, and the outside agencies could rotate on a four-year cycle as Chair of the Executive Committee.

Annual Work Program. OLO recommends establishing an annual work program for the Data Innovation Committee because research shows that jurisdictions that have begun to use advanced data analytics typically begin with a question or an issue and draw on available data in new and different ways to identify an answer or solution.

The Executive Committee could establish the annual work program based on input solicited from:

- The County Council,
- The County Executive,
- The Office of Legislative Oversight,
- Montgomery County Public Schools (MCPS),
- Montgomery College,
- The Montgomery County Planning Department, and
- Data Innovation Committee members.

While MCPS, Montgomery College, and the Montgomery County Planning Department each currently use data analytics and predictive analysis to further aspects of their respective missions, they may identify agency-specific projects that they consider appropriate to include on the Data Innovation Committee's work program. To the extent that they do, the Committee Chair and Vice Chair and the Executive Committee could work with agency staff to define the scope of a project and to identify any agency-specific data needed.

Selection of Committee Members. The Council could select the Data Innovation Committee Chair and Vice Chair by receiving suggestions from the representatives from OLO, the County Council, and the Office of the County Executive on the Executive Committee – with the Chair and Vice Chair subject to approval by the County Council. Subsequently, the full Executive Committee could recommend the remaining Committee members to the Council, who could also be subject to approval. The Data Innovation Committee members could serve a two-year term, with the possibility of reappointment to the Committee by the Council.

Other Considerations. The project description for this report in OLO's FY14 work program states that "OLO will put emphasis on the importance of ensuring that members of the advisory board are interested in the neutral examination of data and not in promoting a particular agenda item or idea." The idea behind the Data Innovation Committee is that the examination of data can help an organization better understand and analyze issues related to its mission.

If the Council establishes a Data Innovation Committee, OLO strongly recommends that the Council take steps to ensure that it remain free from both internal and external political pressures. For example, the Council should:

- Require that Committee members not have a political or ideological “agenda” that they want to advance through the Committee.
- Mandate that agency-specific projects on the annual work program come directly from agency recommendations – to avoid any perception that work program projects are motivated by political considerations or are intended as “fishing expeditions” into agency administration.

The success and usefulness of a Data Innovation Committee will depend on agency cooperation and access to agency data. A perception that the Committee may be driven by political motivations could hinder agency willingness to participate in the Committee or to provide non-public data, if needed, for the Committee’s analysis.

Data Privacy. Data that may be relevant to a Data Innovation Committee project may be subject to various federal, state, or local privacy laws. OLO recommends that the Data Innovation Committee work with the Office of the County Attorney to identify project-specific privacy issues and take steps to ensure compliance with privacy laws.

Timing of Council Consideration

In December 2012, the Council enacted Bill 23-12 on open data. The legislation requires the County Government’s Chief Administrative Officer to develop an Open Data Implementation Plan within 18 months that describes the public datasets under the control of County Government departments and offices and that prioritizes the publication of the datasets on the County Government’s open data portal. Executive Branch representatives reported to the Government Operations and Fiscal Policy Committee in September 2013 that they intend to complete the Open Data Implementation Plan by the summer of 2014.

In order to have a fuller understanding of available County Government data, OLO recommends that Committee and Council formation of a Data Innovation Committee coincide with discussion of the Open Data Implementation Plan when the Chief Administrative Officer releases it next summer.

The table on the next page summarizes the structure described for the Data Innovation Committee.

Table 4-1. Data Innovation Committee Structure

Components of Committee	Options
Number of Committee Members	10-15
Examples of Representation on Committee	Data experts from (examples only): Montgomery College Montgomery County Public Schools Montgomery County Planning Department The University of Maryland The Universities at Shady Grove Johns Hopkins University Non-Profit organizations Private sector businesses using data analytics The Federal Government
Executive Committee	Office of Legislative Oversight representative County Council representative County Executive representative(s) Representative rotating between MCPS, Montgomery College, and the Planning Department Data Innovation Committee Chair Data Innovation Committee Vice Chair
Executive Committee Responsibilities	Development of Work Program Oversight of Committee work/progress Recommendation to Council of Committee Chair and Vice Chair
Committee Work Program Input	OLO County Council County Executive MCPS Montgomery College Montgomery County Planning Department
Selection of Committee Chair and Vice Chair	Recommended by the Executive Committee representatives from OLO, the County Council, and the Office of the County Executive; approved by the Council
Selection of Committee Members	Recommended by the Executive Committee; approved by the Council
Terms of Service	2 years, with the possibility of reappointment by the Council
Data Privacy	Ensure compliance with federal, state, and local privacy laws

Chapter V. Examples of Data

Today, the public has access to thousands of relevant data sources through the federal government, the State of Maryland, and other local jurisdictions, among others. The use of advanced data analytics provides governments previously unavailable opportunities to extract and connect information from a variety of sources. This chapter briefly summarizes examples of data sources and types of data available to the Montgomery County Data Innovation Committee.

A. Federal Government

The federal government releases data through its open data website, www.data.gov. As of November 5, 2013, the website provided access to 94,150 datasets. The information in the table below summarizes the data found in a few of these datasets.

Table 5-1. Examples of Federal Datasets

Data Type and Source	Description
U.S. Census Data <i>U.S. Census Bureau, U.S. Department of Commerce</i>	This dataset provides information both at the county and city level on data points such as total population, population broken down by age and race, housing values, home ownership rates, employment, number of businesses, and retail sales.
Pesticide Residue in Food and Water <i>USDA Pesticide Data Program, U.S. Department of Agriculture</i>	This dataset provides national data on pesticide residues in food and water, with an emphasis on foods consumed by infants and children. PDP data are used primarily by EPA to prepare realistic pesticide dietary exposures for pesticide registration activities.
Climate Data <i>NOAA National Climatic Data Center, National Oceanic and Atmospheric Administration</i>	These data provide North American temperature and precipitation data and is searchable by time period and by location (down to the zip code level).
Food and Drug Administration Recalls <i>U.S. Department of Health and Human Services</i>	These data list manufacturer recalls of food and drug products.

B. State of Maryland

The State of Maryland releases data through its open data website, <https://data.maryland.gov/>. As of November 5, 2013, the State of Maryland's open data website included 169 datasets. The information in the table below summarizes the data found in a few of these datasets.

Table 5-2. Examples of State of Maryland Datasets

Data Type and Source	Description
FY12 payments to vendors <i>State of Maryland Department of Budget and Management</i>	This dataset summarizes payments to vendors that received \$25,000 or more in a fiscal year.
Certified Cover Crops Planted in the Chesapeake Bay Watershed (Annually) <i>State of Maryland Department of Agriculture</i>	This dataset contains information on the number of acres of certified crops planted as a means to determine the overall health of the Chesapeake Bay Watershed.
MDTA Accidents <i>Maryland Transportation Authority</i>	This dataset provides information on accidents at or near a MD Transportation Authority (MDTA) facilities that MDTA police responded to. In Montgomery County, the Intercounty Connector is an MDTA facility.
Public School Pre-Kindergarten Enrollment <i>Maryland State Department of Education</i>	These data list the number of children, by county, enrolled in public school pre-kindergarten programs between 2009 and 2011.

C. Montgomery County

The County Government releases data through its open data website, <https://data.montgomerycountymd.gov/>. As of November 5, 2013, dataMontgomery included 34 datasets. The information in the table below summarizes the data found in a few of these datasets.

Table 5-3. Examples of Montgomery County Datasets

Data Type and Source	Description
Inspections of Retail Food Establishments <i>DHHS Licensure & Regulatory Services Program</i>	This dataset includes information on dates of inspection, inspection results, whether food came from approved sources and was protected from contamination, whether food was cooked properly, whether an establishment has proper sewage disposal, and presences of rodents, insects, toxic substances and pesticides.
Residential Building Permits Issued since 2000 <i>Department of Permitting Services</i>	This dataset includes residential building permits issued since 2000, including street address, data issued and closed, a description of the work, and a declared value for the proposed work.
Open and Closed 311 Service Requests <i>MC311</i>	These datasets list the number of service requests opened and closed by MC311, disaggregated by groups such as topic and department.
2012 Cable Complaints <i>Montgomery County Cable Office</i>	This dataset lists complaints filed with the Montgomery County Cable Office since January 2010 and includes information on the cable provider, the complaint issue, and the location associated with the complaint.
Percentage of Students Receiving Free and Reduced-Price Meals (FARMS) <i>Montgomery County Public Schools</i>	These data provide information on the percentage of students in each school who participate in the National School Lunch Program, disaggregated by gender, race, and ethnicity.
Facility Data <i>Montgomery County Public Schools</i>	These data describe the age of MCPS school facilities, when they were last renovated/modernized, the acres of land where the facilities are located, and whether other programs are jointly-housed at a facility.
Capital Budget Expenditure Schedule <i>Montgomery College</i>	These data summarize projected capital spending for six years, including cost of land, site improvements and utilities, and construction.
Safety and Security Annual Security Report <i>Montgomery College</i>	Among other information, this report includes annual data on criminal activity on the College's three campuses.
Montgomery County Development Pipeline <i>Montgomery County Planning Department</i>	These data include the number of dwelling units approved by the Planning Board, by project and master plan area.
Population Forecasts <i>Montgomery County Planning Department</i>	These data include projections on population growth in the County by area through 2040.

D. Other Local Jurisdictions

Some other jurisdictions, such as the District of Columbia, release data through an open data portal. Other jurisdictions, such as Fairfax County, VA and Frederick County, MD provide some datasets on department websites. The information in the table below summarizes examples of datasets from other local jurisdictions.

Table 5-4. Examples of Datasets from Other Local Jurisdictions

Data Type and Source	Description
<p>Accepted Zoning Applications <i>Fairfax County, VA Department of Planning and Zoning</i></p>	<p>These data break down the number of acres used for different purposes, such as commercial, industrial, and residential, in Fairfax County.</p>
<p>Available Regulated Child Care Programs <i>Fairfax County, VA Office for Children</i></p>	<p>This database allows users to search for available child care options that are licensed by Fairfax County and/or the Commonwealth of Virginia.</p>
<p>Procurement and Contracting Awards <i>Frederick County, MD Finance Division</i></p>	<p>These data list all contracting awards made by the Frederick County Government.</p>
<p>Purchase Card Transactions <i>D.C. Office of Contracting and Procurement</i></p>	<p>This dataset includes Purchase Card transactions made by City employees.</p>

Chapter VI. Agency Comments

The Office of Legislative Oversight circulated a final draft of this report to the County Government, Montgomery County Public Schools, Montgomery College, and the Montgomery County Planning Department. OLO appreciates the time taken by agency representatives to review the draft report and provide technical feedback.

Written comments on the final draft report from the Chief Administrative Officer begin on the next page. Any written comments from the agencies will be included in the packet for Committee review of the report. MCPS, Montgomery College, and the Planning Department have all indicated their intent to participate in the work of the Data Innovation Committee.



OFFICE OF THE COUNTY EXECUTIVE

Isiah Leggett
County Executive

Timothy L. Firestine
Chief Administrative Officer

MEMORANDUM

November 25, 2013

TO: Chris Cihlar, Director, Office of Legislative Oversight

FROM: Timothy L. Firestine, Chief Administrative Officer *Timothy L. Firestine*

SUBJECT: Draft Report - Montgomery County Data Innovation Committee

I am in receipt of your draft report dated November 6, 2013, recommending a new Advisory Board identified as a "Montgomery County Data Innovation Committee." The report's recommendation is in response to the Council's request for OLO to examine the feasibility of creating an advisory board with the expertise to conduct analytics on large scale, complex data ("big data").

We appreciate your efforts in examining the potential methods by which Montgomery County could strengthen its efforts around "big data" by tapping into the pool of County residents and businesses with a data science background.

In general, having a pool of talented County residents and businesses could provide useful contributions to data analytics within County agencies. On the other hand, a formal committee approach is not consistent with the main purpose of the "open data" concept which is to democratize the use of both structured and unstructured information/data, and provide it to individuals or groups who would make use of it for the public good.

As to the structure of OLO's recommended Data Innovation Committee, the proposed Executive Committee does not include permanent representatives from all County agencies (MCPS, MNCPPC, WSSC, Montgomery College, HOC), municipalities, and the State, which could limit its effectiveness/influence.

Also, I would like to highlight the following points:

- Since the referenced work program for this proposed cross-sector cross-agency committee seems to be the key to its usefulness, there needs to be defined guiding principles, unified desired outcomes, and dedicated resources for this group. Although we cannot speak for the other County agencies, currently the appropriate resources for staffing this committee are not

available and an adequate inventory of meaningful and easily accessible/analyzable datasets is not yet available.

- If the purpose of this committee is to provide a skill set(s) not currently available in County government, it would be helpful if the skill set(s) could be identified. CountyStat is evolving to address the opportunities and challenges of “big data” and it is entirely possible to add the necessary skill set(s) and/or tools to its current process.

We recommend that the Council postpone considering this recommendation until the County’s Open Data Implementation Plan has been in place for at least one year and aspects of it have been effectively integrated with our existing systems such as ERP, MC311, and other departmental service-focused databases/systems and social media tools.

Thank you for the opportunity to respond to this draft. If you have any questions, please feel free to contact Assistant Chief Administrative Officer Fariba Kassiri.

TLF:dh

cc: Fariba Kassiri, Assistant Chief Administrative Officer
Sonny Segal, Director, Department of Technology Services
David Gottesman, Manager CountyStat
Dan Hoffman, Chief Innovation Officer
Dieter Klinger, Chief Operating Officer, Department of Technology Services



Appendix

Appendix A

The information below describes the missions of the ad hoc groups cited in Chapter IV that have been developed at the federal, state, and local levels for the purpose of pursuing innovation in government.

President’s Council of Advisors on Science and Technology¹

“On April 27, 2009, President Obama announced the President's Council of Advisors on Science and Technology (PCAST). PCAST is an advisory group of the nation’s leading scientists and engineers who directly advise the President and the Executive Office of the President. PCAST makes policy recommendations in the many areas where understanding of science, technology, and innovation is key to strengthening our economy and forming policy that works for the American people.

“Beginning in 1933 with President Franklin D. Roosevelt's Science Advisory Board, each President has established an advisory committee of scientists, engineers, and health professionals. Although the name of the advisory boards have varied over the years, the purpose of each remains the same—to provide scientific and technical advice to the President of the United States.”

Federal CIO Council’s Innovation Committee²

“To enable the transformation to a 21st Century Government that serves the American people more effectively, agencies have been innovating with less and strategically investing in information technology (IT). The Innovation Committee focuses on relevant topics such as the use of modern technologies to deliver digital services to citizens and businesses, deployment of mobile technology within Government, modular IT development strategies, and using Federal data as strategic resource to enable agency mission delivery and grow the economy.”

Federal Leading Education by Advancing Digital (LEAD) Commission³

“Answering a challenge from the Federal Communications Commission (FCC) and the U.S. Department of Education, experts from across the education and technology space formed the Leading Education by Advancing Digital (LEAD) Commission. The Commission has developed a blueprint, detailing the opportunity for using technology as a catalyst to transform and improve American education. The LEAD Commission has incorporated input from a cross-section of teachers, parents, local government officials, school officials, students and education technology industry leaders to develop its findings and guide the creation of the blueprint.”

¹ <http://www.whitehouse.gov/administration/eop/ostp/pcast/about>

² <https://cio.gov/about/groups/innovation-committee/>

³ <http://www.leadcommission.org/about-lead>

New Hampshire Governor's Commission on Innovation, Efficiency & Transparency⁴

"The commission brings together experts from the private business, nonprofit, higher education and local government sectors to make recommendations for modernizing state government, improving efficiency, measuring the performance of state agencies and improving transparency to citizens....

"The Governor's commission includes members with relevant experience in areas including performance measurement, reorganizations, efficiency and operational improvements. Members will examine the organization of state government and opportunities for reducing redundancies or streamlining services; how to improve government processes, including the use of technology; the state's contracting and relationship with its nonprofit partners, quasi-governmental agencies and local communities; performance metrics for state government and for improving transparency; and other methods that members believe will improve the efficiency and effectiveness of state government in order to deliver cost-effective quality services to the people of New Hampshire."

Kansas City, MO Mayor's Challenge Cabinet⁵

Mayor Sly James [appointed] a cross-sector group of young professionals to serve as a creative advisory panel, with the goal of formulating forward-thinking civic policies that reflect the needs and wants of emerging leaders. Dubbed the Mayor's Challenge Cabinet, the group will work directly with the City's recently hired Chief Innovation Officer and the Code for America fellows.... The goal is to fundamentally transform the way City Hall approaches problems, working strategically to develop efficient, effective, and enhanced municipal operations.

⁴ <http://www.governor.nh.gov/media/news/2013/pr-2013-05-29-innovation-commissio.htm>

⁵ <http://kcmayor.org/challenge-cabinet>