6. Opening Data to Spark Innovation

First published as a blog post on March 23, 2017

CAP Goal Statement: *Fuel entrepreneurship and innovation, and improve government efficiency and effectiveness by unlocking the value of government data; adopting management approaches that promote interoperability and openness of these data.*

Data can support the quality of important choices citizens make every day. For example, the federal government's Open Data initiative has led to the creation of a useful one-stop resource: the College Scorecard, hosted by the US Department of Education. And, the Open Data initiative has led to similar information resources in other parts of the economy.

Background. Access to government information has exploded over the past two decades. For example, during the expansion of the Internet in the late 1990s and early 2000s, the federal government created a one-stop directory of all federal websites. But that didn't necessarily provide access to agency data sets.

One of the distinctive initiatives of the Obama Administration was its advocacy of Open Government. A key element of this broad initiative was to make federally-created data even more openly available to the public and entrepreneurial businesses. It claimed that open data "strengthens our democracy and promote efficiency and effectiveness" as well as increasing transparency and accountability. But more concretely, ready access to government data can empower individuals and businesses to make better decisions. Its first tangible step in 2009 was to create a one-stop website, Data.gov, where agencies could register their "high value" data sets for public access.

President Obama issued an executive order in 2013 that formalized and expanded this effort. The order requires agencies to make government data "open and machine-readable" wherever possible. This was supplemented with a directive from the Office of Management and Budget (OMB) to "manage information as a strategic asset" and make it openly available—while ensuring privacy, confidentiality and national security.

Objectives of the Cross-Agency Priority Goal. In early 2014, the Open Data initiative was designated as one of the federal government's 15 cross-agency priority (CAP) goals. The Open Data goal has two primary objectives:

- ***Fuel economic growth and innovation through the increased use of federal Open Data.*** This includes the release of additional "high priority" datasets identified by the agencies and/or the public. It also includes the development of tools and support for agencies and data users, as well as sponsorship of quarterly events to encourage the use of agency data, such as "data jams," code-a-thons and incentive prizes.

- ***Make open and machine-readable "the new default" for all government information.*** This includes the development of "enterprise data inventories" across all agencies and making this data "discoverable" by the public.

At the same time that the Open Data initiative was designated a CAP Goal, the Data.gov website was relaunched, with a simpler user interface that makes it easier to search for relevant data sets.

Governance Structure. This CAP Goal has three co-leads:

- the U.S. chief information officer, located in OMB;

- the U.S. chief technology officer (CTO), located in the Office of Science and Technology Policy in the White House and

- the deputy secretary of the Department of Transportation.

They were supported by a small staff in the General Services Administration (GSA), whose primary responsibility is the operation of the data.gov website. The staff also helps in the coordination and standard-setting that occurs across a broader community, the Federal Interagency Open Data Working Group. This working group is co-led by the three CAP Goal leaders and it meets biweekly. The group also holds an open quarterly meeting to directly engage with civil society groups. In addition, daily communications among the 800 government members takes place on GSA's listserv.

Evolution Over Time. From the time of its launch in 2009 until 2013, Data.gov was a place where agencies could register their data sets for public access. In conjunction with the 2013 Executive Order, OMB launched a Project Open Data Dashboard to track agency progress toward implementing the executive order on a quarterly basis.

In 2013, under the leadership of U.S. CTO Todd Park and U.S. Deputy CTO Nick Sinai, the Obama Administration began using term-limited Presidential Innovation Fellows to help agencies unlock data, engage external stakeholders and generally focus on the use of Open Data.

Along with the 2014 relaunch of Data.gov, the Obama Administration released an Open Data Action Plan, recommitting the Administration to its policies, in the context of the broader National Action Plan for Open Government—an international commitment to Open Government that is revised periodically with new commitments (most recently in 2015).

In 2015, the White House appointed DJ Patil to serve as the federal government's first Chief Data Scientist. As a part of the office of the Chief Technology Officer, he helped develop data policies but also helped "foster partnerships to help responsibly maximize the nation's return on its investment in data, and help to recruit and retain the best minds in data science to join us in serving the public," according to a White House statement.

During 2016, the Office of Science and Technology Policy hosted a series of Open Data round-tables, including a White House Open Data Innovation Summit in late 2016. This summit crystalized the energy and enthusiasm for the movement among public and private sector trailblazers and entrepreneurs committed to using data to improve government efficiency, citizen health and wellness and economic opportunities.

In addition to the CAP Goal, the topic of Open Data has gained an increasing circle of interest outside of the federal executive branch, since the broader movement encourages opening private sector, non-profit, state and local data. For example, there is a corporate advocacy group, the Data Coalition, and there is congressional interest in legislating Open Data initiatives. In the last Congress, the OPEN Data Act—which would require agencies to standardize their data in an open format that is machine-readable—cleared the Senate and is expected to be reintroduced in the current Congress.

Progress to Date. Agencies have opened many of their data sets to the public. As of mid-March, 195,016 have been cataloged on Data.gov. Agencies are also beginning to document their data sets in a common format, which improves the chances that researchers and the

public can find them. Interestingly, over 50 states and localities have adopted the federal open data standards and are posting their own data on Data.gov so they can be integrated with federal agency data.

Examples of Uses of Open Data. The emphasis of the Open Data initiative has evolved over time from the supply of data, such as via Data.gov, to the use of these data. Agencies were asked by the Office of Science and Technology Policy to actively find out how their newly-opened data were being used. Some uses were internal to the government. Others were by private sector entrepreneurs. For example:

- **College Scorecard**—The Department of Education launched a redesigned tool in 2015 to help parents and student to compare college costs and outcomes. According to the Department, the "refresh includes more than 1,700 data points for more than 7,000 institutions of higher education." It also includes: "statistics on debt, federal loan repayment, completion rates and post-college earnings of alumni in an easy-to-understand format." In addition, it offers a streamlined application for federal student aid.

- **The Opportunity Project.** Hosted by the Census Bureau, this project set out in early 2016 to create data sets and digital tools to help communities navigate information about quality housing, schools, jobs and transportation. The project has worked with cohorts of tech companies, non-profits and federal-state-local agencies to create tools to help connect unemployed with skills and jobs. For example, one partner, LiveStories.com, worked with Marin County, CA, health officials to bring together stakeholders around a shared goal of reducing obesity in their community.

- **Police Data Initiative.** This initiative, launched in 2015, is intended to increase transparency and accountability. A year later, more than 50 police departments had joined, releasing over 90 data sets on police-citizen interactions (use of force, traffic stops, pedestrian stops, etc.). The data sets that became available via the Open Data initiative became the basis for a broader Public Safety Data Portal sponsored by the non-profit Police Foundation.

- **Zillow.com,** a private sector on-line nationwide real estate media company, is largely based on Open Data from federal, state and local sources. It provides estimates of home and rental values to allow consumers to make more informed choices. It draws on federal census data, labor statistics and housing data.

Next Steps. The various federal data communities continue to meet on a periodic basis, providing additional data sets as well as refreshing examples of how these data are being used. While in early 2017, media reported that several high-profile data sets had been removed from federal websites, nearly 2,000 new data sets have quietly been added to data.gov as of mid-2017.

Unlike any of the other CAP Goals, the Open Data goal has an interesting international dimension, in that it can be seen in the context of a global "megatrend." This trend is powered in large part by the belief that it can contribute to new economic growth. A World Bank study in 2014 cites economic opportunities in the range of $3–5 Trillion a year. In addition, there is a Global Open Data Index that assesses the extent to which governments around the world are open with their data, as judged by citizens, not by the governments themselves.

As a result, the US Open Data initiative—even beyond the CAP Goal designation—can learn as well as contribute to best practices among peers across the world, in addition to practices at the state and local levels.