CAPITALIZING ON THE
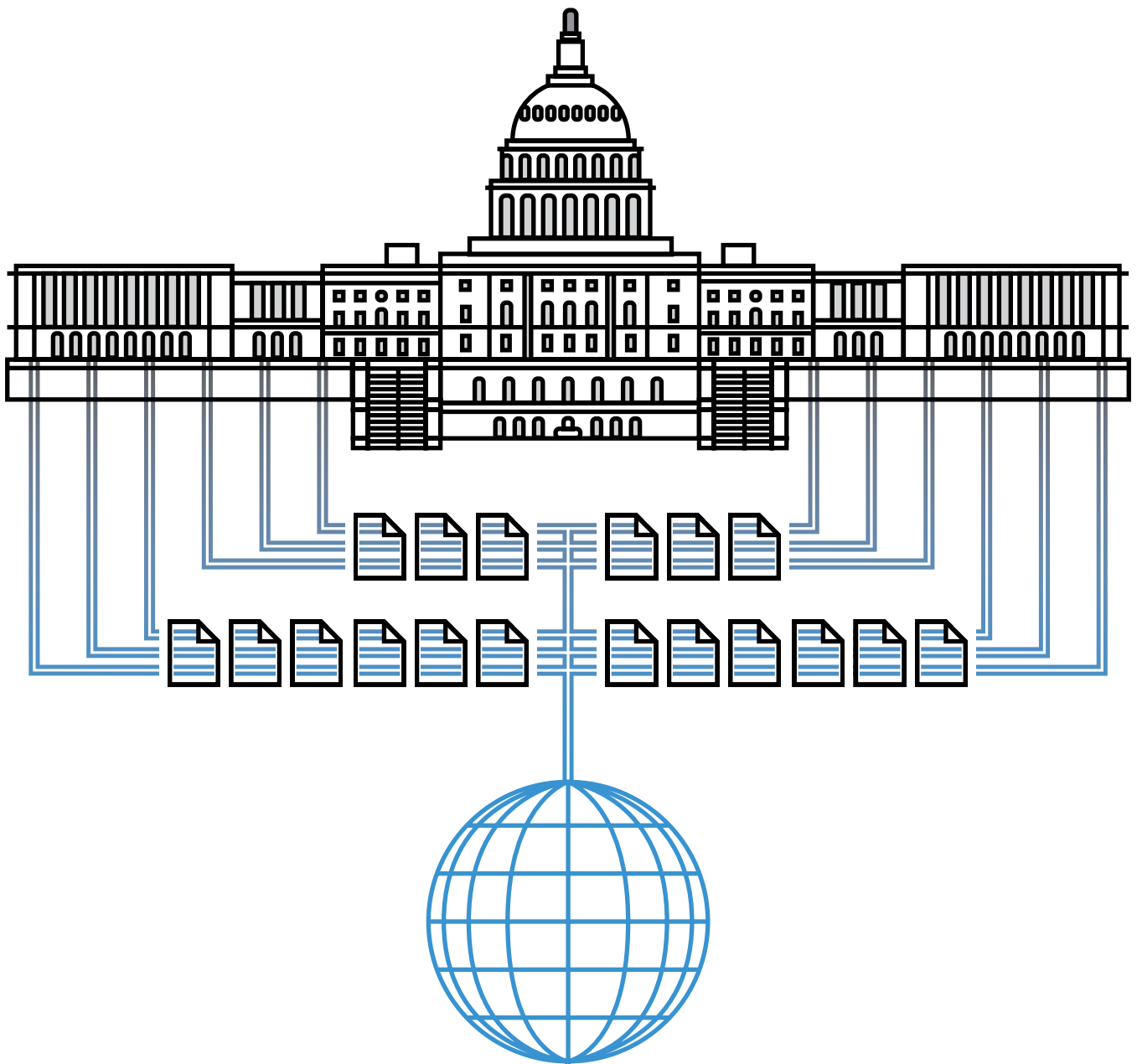# OPEN DATA REVOLUTION

EMC²
BUSINESS
PARTNER

REDEFINE FEDERAL IT

EMC Federal Summit 2014
October 22, 2014 • Washington, D.C.

Click here to save the date!

Sponsored By:

carahsoft.

EMC² **Pivotal** **RSA** **vm**ware®

EMC's Audie Hittle, federal chief technology officer, Isilon Storage Division, EMC Corporation, tells us how open data is transforming our world

Government has entered a new era, and it's being powered by open data. Open data has been driving us to strive for a more transparent, participatory and collaborative government, and it helps us deal with the complex challenges we face today.

But how will the data-driven era shape the future of government?

It's simple. With open data programs, governments no longer have to be the sole provider of services. Government data can now serve as a platform for developers and citizens to access and create new services. As government spurs innovation, businesses and services will continue to sprout, many of which government workers could never imagine, let alone have the resources to create.

Take the findings of a recent report by the Department of Commerce's Economics and Statistics Administration, "Fostering Innovation, Creating Jobs, Driving Better Decisions: The Value of Government Data." It states that:

- Government data guides trillions of dollars of investments each year by local, state and federal governments; households; institutions; and private companies.

- Firms that intensively use data from Commerce's statistical agencies often combine it with other government and private-sector data to create $24 billion to $221 billion in annual revenues.

- Government data is uniquely comprehensive, consistent, credible, relevant and accessible, and still protects confidentiality.

The report represents only a portion of the power of government data to drive markets and improve government efficiency.

We believe that government must embrace this new era of data innovation. When it does, officials will be able to tackle some of the most complex public-sector challenges and create a brighter, more promising future for citizens.

CONTENTS

EXECUTIVE SUMMARY

For years, data advocates have championed open data initiatives to create more efficient government and economic opportunities. They argued that through open data, government could support emerging markets rather than create additional expensive public programs. And today, that vision is coming to life.

Take the case of Trulia, an online residential real estate website for homebuyers, sellers, renters and real estate professionals. Trulia has a feature that creates a visualization to allow users to calculate commute times and view property values, home values, rental prices, schools, points of interest, natural hazards and crime data all in one spot and all based on public data. To create this service, Trulia pulled data from the U.S. Geological Survey and the Federal Emergency Management Agency. Trulia is a good example of how government can create information ecosystems that facilitate innovation across all sectors.

The open data landscape is quickly changing, and one way agencies can be ready for the changes is to learn best practices from public-sector open data pioneers. That's why in this report we explore the groundbreaking work of the Food and Drug Administration, Commerce and Data.gov. In this guide, you'll find case studies on how each agency has taken advantage of open data to power a variety of initiatives:

- The Food and Drug Administration's openFDA initiative makes FDA's data publicly available and accessible in structured, computer-readable formats that make it easy for developers to build mobile and web applications and data visualizations for researchers to quickly pull FDA data through application programming interfaces. APIs are used to create a set of definitions that dictate the ways one piece of computer software communicates with another. You'll learn from Sean Herron, a former Presidential Innovation Fellow who led the project, about the power of working with your user community to be sure that the data released is useful and valuable.

- Department of Commerce's International Trade Administration recently launched a Developer Portal called developer.trade.gov. This site gives software developers access to data on U.S. exports and international trade through APIs. In this case study, Mike Kruger, director of digital engagement at Commerce, will provide you with an understanding of how to collaborate on open data initiatives across an agency and with industry.

- EMC Isilon's chief technology officer for the federal market, Audie Hittle, will explain technology's role in powering open data initiatives. He provides intriguing case studies from the Centers for Medicare and Medicaid and Walter Reed National Military Medical Center and the work that EMC has done to support open data programs government-wide.

- The Data.gov program continues to power innovations inside and outside government. In our final case study, you'll learn from Jeanne Holm, Data.gov evangelist, about the outlook of open data. She also gives an update on Data.gov's transformative work.

From Trulia to our government case studies, it's clear that open data has reached a tipping point – a fact we also see from our GovLoop survey of 208 public-sector

"DATA SHOULD BE FREELY SHARED WITH THOSE WHO CAN BENEFIT FROM IT TO IMPROVE EFFICIENCY, PRODUCTIVITY, COLLABORATION, RESPONSIVENESS, ACCOUNTABILITY, WHILE DECREASING COST AND MINIMIZING THE OBSTACLE IN ACHIEVING DESIRED RESULT."

Govloop survey respondent

workers. The survey found that more than 50 percent of respondents believe that open data collaboration is mission-critical or very important. Additionally, 51 percent believe that transparency and accountability are driving their agency to adopt open data practices.

"Data should be freely shared with those who can benefit from it to improve efficiency, productivity, collaboration, responsiveness, accountability, while decreasing cost and minimizing the obstacle in achieving desired result," a survey respondent said.

Today, government officials must think about how data can help them make decisions that uplift communities, empower citizens to take action, and create more efficient and effective agencies. Whether it is Commerce, Data.gov or FDA, we are now witnessing a philosophical shift in how public-sector officials think about their high-value and authoritative data. Our guide will help you to:

- Understand how to drive action from your open data programs.

- Identify ways in which government has excelled with open data.

- Provide best practices to prepare your agency to adopt open data.

- Promote case studies from leading agencies adopting open data.

Welcome to a new era, where government is an enabler of new services, and not just the provider. Now's the time to learn how open data can empower your agency.

50%
BELIEVE THAT OPEN DATA COLLABORATION IS MISSION-CRITICAL OR VERY IMPORTANT

51%
BELIEVE THAT TRANSPARENCY AND ACCOUNTABILITY ARE DRIVING THEIR AGENCY TO ADOPT OPEN DATA PRACTICES

MedWatcher is a free mobile and web app that allows patients to understand and report side effects of drugs, medical devices and vaccines. The service was created in 2010 and is an early example of the power of using FDA data to build new services for patients. Its use of FDA's adverse drug events data and information collected from the community informs more people about side effects from medications and improve FDA's ability to collect and disseminate such data. MedWatcher is an example of how open data can support new services, which government may not have the resources to create.

For years, FDA has been collecting and sharing data related to adverse drug events. But until recently, the agency did not have an efficient way to share data, make it easily accessible or enable users to contribute to it.

Previously, companies and the public often had to submit Freedom of Information Act requests to receive data. In some cases, the data was already public, but even then, it was often hard for developers to find accurate and timely data on which to build applications. FDA found that for companies to truly make the most of its data, it needed to make the data accessible.

That's why FDA created openFDA. With openFDA, data is now publicly available in structured, computer-readable formats. This makes it easy for developers to build mobile applications and web services and create data visualizations. Researchers can also use openFDA to quickly access FDA data through APIs. Now developers can easily integrate the data into other sources.

"OpenFDA really shows the power of having a very focused effort to make sure that data is usable and accessible to citizens. It's incredible to see what community members can create when provided access to data, which in many cases were things that we never imagined," said Sean Herron, a former Presidential Innovation Fellow who served on the team that created the openFDA platform.

DATA DISCOVERY: THE PROCESS OF DELIVERING VALUE

To be sure they were delivering value to their data consumers, FDA officials first needed to know what kinds of data would be most useful for their users. They started by conducting a thorough inventory of their data, exploring which datasets were most actively used and already released.

They found that although they had released a considerable amount of data, many groups that could find the data valuable — startups, developers or citizens — rarely used it. Officials assumed this was because those without a deep knowledge of the FDA did not know the data existed, or because there was little documentation and structure around the information. Herron hoped to change that.

After taking inventory of the data, the openFDA team identified three of the most valuable datasets and decided to build a platform around them. The three core datasets were:

- Adverse drug event report: An adverse event is submitted to FDA whenever an undesirable experience is associated with use of a drug. This includes side effects, product use errors and quality or therapeutic failures. Since 2004, FDA has been collecting this data and has more than 4 million records. It is important to note that the reporting of adverse drug events are completely voluntary in the United States, but sharing information can help combat unintended consequences of medications.

- Product and food recalls: This openFDA API is for all product recalls FDA monitors. At times, the agency may remove a product from the market to protect the public. Companies almost always do recalls voluntarily. FDA's role during recalls is to oversee strategy and classify the recalled product according to the level of hazard. FDA then posts all recall information in its Enforcement Reports.

- Product labels: FDA also collects data around product labels, such as those on the back of pain medication bottles listing active ingredients and dosage. This data is useful for researchers, citizens and developers to help bring more clarity to what is in medications.

FDA officials were determined to maximize the utility of these three datasets. "We really wanted to deep dive into those three datasets and provide a great experience for developers wanting to use them," Herron said.

FDA conducted user testing, did small beta releases and iteratively built openFDA as the team gained more insights from their community. But doing so did not come without challenges.

openFDA provides open APIs, raw data downloads, documentation and examples, and a developer community for an important collection of FDA public datasets.
About openFDA »

ADDRESSING DATA CHALLENGES WITH COLLABORATION

"Although computing resources are cheaper, faster and more effective than they have ever been in the past, there are still many challenges to make any kind of dataset available on the Internet," Herron said. "You need to have a skilled group of people from across the spectrum to ensure you are ensuring security and privacy, providing the data in a usable interface, and building a solution that can scale to meet demand."

To work through the final challenges on preparing data to be public, FDA brought together a team of software engineers, user interface professionals, and policy and legal experts.

The openFDA initiative's beta version was publicly re-leased in June 2014. Since then, the agency has continued

to improve and build on it. Within the first week of the beta release, it had more than 1 million queries.

"We saw a number of applications and websites created that allow FDA information to be mashed up with other datasets, making information easier for citizens to under-stand," Herron said.

OpenFDA removed the previous barriers to FDA data. Rather than having to download gigabytes of data with sometimes confusing XML formatting, openFDA now allows easy developer access via a JSON-based API. It is easy for software developers to use FDA data to create their own applications and websites, outside the scope of the agency.

OPEN DATA LESSON: THE USER'S NEEDS COME FIRST

The major lesson learned from FDA is that in order to excel with open data, organizations must have a firm grasp on user needs. By working so closely with its users, FDA was able to figure out what the three most high-val-ue datasets were and deliver the APIs and resources that were in demand.

"My users are the citizens of the United States of America, so I need to be responsive to their needs," Herron said.

Herron is not alone in his transformative work. In our next section, we explore the open data work at Com-merce, which is helping "America's Data Agency" leverage information in new ways.

LEARNING FROM AMERICA'S DATA AGENCY

Since the founding of the United States, Commerce has been collecting important data about the American economy. From the first Census in 1790 to the founding of the Patent Office in 1802, the department has been America's leading data collection agency.

"We are America's Data Agency both by mission and mandate," said Mike Kruger, Director of Digital Engagement at Commerce.

Department-wide, agencies are collecting data about nearly every kind of economic indicator – and many of them are trying to make the data useful. Nowhere is this better illustrated than at Commerce's National Oceanic and Atmospheric Administration.

NOAA collects more than 20 terabytes of data per day, nearly twice that of the entire printed collection of the Library of Congress. The data encompasses everything from Doppler radar systems to weather satellites, from tide gauges to information about ships and aircraft. But right now, only a fraction of that data is available in a timely manner to the public.

The NOAA case is an example of one of the main challenges facing the agency: quantity over quality. Although the agency collects troves of data, NOAA leaders wanted to be sure that the data released was actually useful for developers. To understand how to optimize their publicly accessible data, NOAA released in February 2014 a request for information (RFI) to engage with the private sector.

The RFI will help NOAA understand how to make its data more accessible to the public. By opening its data and increasing engagement with developers, NOAA's community and the private sector, department officials hope to turn those 20 terabytes of data into meaningful public-sector solutions.

NOAA received more than 70 responses to the RFI and is working with respondents and other agencies to formu-late a plan for making the data more useful. Ultimately, NOAA's goal is to create a data ecosystem that will be adaptable as the public's data demands change.

NOAA's efforts to find new ways to open and use its data are part of a larger departmentwide data strategy at Commerce. NOAA officials recognize that the department holds vital data that can grow the economy, and by making the data freely accessible, they can support citizen activists in creating new public services and help businesses unlock strategies to gain a competitive edge.

In fact, Commerce's 2014-18 strategic plan has a pillar dedicated to data innovation. Through it, Commerce seeks to:

- Maximize the value of Commerce data to governments, businesses and communities.

- Make data easier to find, access and use.

- Partner with other government agencies and the private sector to improve interoperability and dissemination.

- Collect and manage data more efficiently and less intrusively.

- Spark innovation and fuel entrepreneurship with open dialogue and integrated feedback.

Open data will play a critical role in helping the agency meet these goals.

"There are 6 million employer businesses in the U.S., and we obviously can't touch them all, so using our open data with third parties provides the ability to reach additional American businesses and is vital to the ongoing data efforts," Kruger said.

One example of how Commerce is already having success comes from its International Trade Administration (ITA).

"WE WANT TO MAKE OPEN DATA NOT JUST A COMPLIANCE ISSUE, where we are just checking the boxes, BUT INSTEAD PART OF OUR CULTURE AND PART OF WHO WE ARE AND WHAT WE DO."

Mike Kruger, Director of Digital Engagement at Department of Commerce

OPEN GOVERNMENT AT COMMERCE: TRADE DEVELOPER PORTAL

ITA is responsible for strengthening the competitiveness of U.S. business at home and abroad by promoting trade and investments. To meet its objectives, ITA recently launched developer.trade.gov, a portal that gives software developers access to information on U.S. exports and international trade via APIs. The data is designed to help U.S. small-business exporters understand how to expand overseas. Currently the portal includes data on:

- Trade events.

- Market research.

- Trade leads.

- Locations of domestic and international export assistance centers.

- Trade news and articles.

Since launching in July 2014, the initiative has helped advance the agency's data objectives – and ITA has also heard case studies about success from users.

"Within 72 hours of developer.trade.gov being live, we found that a Fortune 20 company who used to literally copy and paste our PDFs into their intranet, now was using our APIs to immediately transfer them," Kruger said. "Now, an employee can go from simply copying and pasting to doing more valuable work for the company, and the company also gets every update as soon as we publish something new."

Within the first two weeks, ITA saw its reach expand. For example, Kruger said one company had indexed all of the ITA market research and upcoming events and started to share that information with thousands of its users, which allowed ITA to reach new markets it hadn't before.

HOW COMMERCE IS CREATING A FRAMEWORK FOR OPEN DATA

To help galvanize open data at Commerce and meet the objectives of its strategic plan, department officials have taken three important steps to improving how it uses data.

1 HIRING A CHIEF DATA OFFICER

Currently, the department is looking to hire a chief data officer. By announcing this role, Commerce follows the lead of the Federal Communications Commission, Federal Reserve Board and National Institutes of Health, all of which were early hirers of CDOs. The CDO will be responsible for managing data departmentwide and thinking of new strategies to spur innovation and collaboration efforts.

"What we don't do, or what we have not done well, is having Census and Patent and Trademark Office or International Trade Administration collaborate and build common systems," Kruger said. "The CDO is going to be doing a lot of that internal work to improve the mechanisms between Commerce, so that we are ultimately better serving the customer."

2 CREATING AN OPEN DATA ADVISORY COMMITTEE

Commerce will also create a data advisory council. Comprised of private-sector leaders, the council will advise the department on how to best leverage data.

"We are going to be running the advisory committee that is going to be full of data users, figuring out how we help move the economy forward by releasing our data," Kruger said.

Commerce has also been hosting open data roundtables. The most recent one was co-hosted by Commerce, New York University's GovLab and the White House Office of Science and Technology. The roundtable brought together 21 companies from various sectors and helped inform a discussion on how Commerce can better use data.

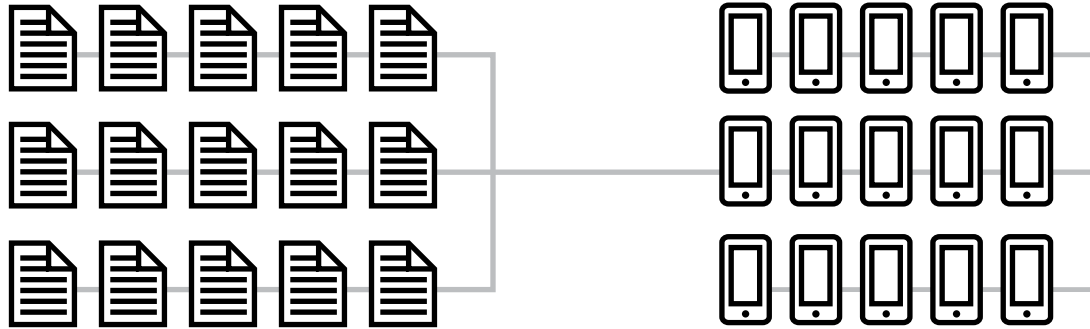### 3 SHOWING THE POWER OF OPEN DATA

The final efforts under way at Commerce are to continue to build a data-driven culture.

"We want to make open data not just a compliance issue, where we are just checking the boxes, but instead part of our culture and part of who we are and what we do," Kruger said.

Commerce has created a model for data-driven innovations. But none of these efforts could have been realized without some success from the groundbreaking work of Data.gov.

Recently, Audie Hittle, federal chief technology officer, Isilon Storage Division, EMC Corporation, sat down with GovLoop to discuss how data can transform the federal government.

GovLoop: How do you see big data and open data impacting government?

Hittle: If our federal government can truly implement a more transparent and open environment that leverages the proven insights and efficiencies available from big data, I'm really optimistic that the impact can be truly transformational. I think we're now just scratching the surface of what is possible if our government continues to invest appropriately in these technologies that can generate such dramatic returns on investment.

As an example, we have implemented EMC's Isilon solutions in certain environments where we've been able to reduce manpower and staffing requirements by around 90 percent due to Isilon's built-in intelligence and automation. That's a tremendous return on investment on manpower and staffing alone, which is one of the highest areas of interest and demand across the federal government.

If we can do that, there's no reason to believe big data's impact will be anything less than transformational.

GovLoop: What do you see as some of the challenges with big data?

Hittle: According to the MeriTalk survey, "The Big Data Cure," – which surveyed 150 federal executives focused on health care and health care research – many agencies are not yet taking steps to prepare for the influx of data. Less than one in five say that their agency is prepared to work with big data, and only roughly a third have invested in IT systems or solutions to optimize data processing or invested in the training of IT professionals to manage and analyze big data.

Meanwhile, new mobile technologies are emerging that are demanding agencies to prepare for the future – like mHealth and machine-to-machine solutions. I'm seeing that agencies are highly interested in these capabilities, but they must also take on the cybersecurity, governance and compliance challenges to fully reap the associated big data benefits.
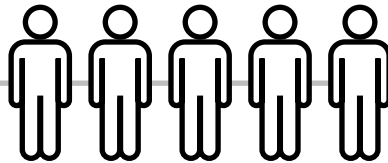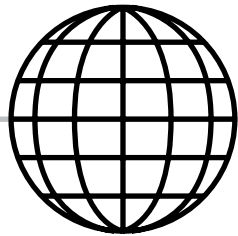
GovLoop: How can EMC help organizations leverage big data and open data?

Hittle: When it comes right down to it, EMC is a truly innovative company with extremely creative people and a vast federated product portfolio. This translates into options and opportunities for customers and partners who are just beginning to explore or preparing to do a deep dive into big data and open data.

EMC offers solutions that range from on-site or remote training and consultations, to fully integrated converged solutions, which can be designed with the client, manufactured and configured off-site, and then delivered, set up and operational in a matter of hours – like our VCE Converged Infrastructure solutions. And, the ultimate in operational flexibility of Isilon is truly a breakthrough when it comes to creating a collaboration environment for any federal agency.

GovLoop: Can you share case studies of work you have done with the public sector?

Hittle: One example from an open data perspective is an initiative called research.data.gov. When I was on active duty in the Air Force, I served on the executive board of the Federal Laboratory Consortium for tech-

nology transfer, and one of the biggest challenges we had then was creating awareness and getting the word out on the technologies and technology trends for opportunities associated with our nation's 700-plus federal laboratories. As of June 2014, under the Open Data Initiative, research.data.gov provides machine-readable data on these 700-plus research and development facilities and gives access to over $100 billion annually in federally funded research, development and manufacturing techniques that may be used by the public, entrepreneurs and innovators to explore, prototype and test new technologies.

From a U.S. citizen's/federal taxpayer's perspective, big data and its associated analytics are already having a major impact. According to the May 2014 White House big data report, the Centers for Medicare and Medicaid Services have begun using predictive analytics software to flag likely instances of reimbursement fraud before claims are made. This is also happening at the Internal Revenue Service, and in the Medicare and Medicaid and health care area, this has reportedly already saved over $115 million in fraudulent payments and is reportedly saving $3 for every $1 invested in the program in the first year alone. Some studies project that such proactive fraud analytics could save taxpayers hundreds of billions of dollars per year and freeing-up funding for other vital initiatives.

GovLoop: **Any additional case studies you could share?**

Hittle: As another example, EMC works with Walter Reed National Military Medical Center in Bethesda, Md. The CIO, Navy Lt. William Walders, is definitely one to watch – a terrific individual, really on the ball.

I had an opportunity to meet him at the May 2014 EMC World and sit in on his presentation – his initiatives and case study examples are incredible. As is the case with a majority of healthcare organizations, his is facing massive big data challenges – working to maximize valuable insights across IT, within budget, in a combined 32 hospitals, with roughly 250 IT staff.

His team manages the local data collection, as well as remote petabytes of active archive data that is being used to store data to support analytics and access to support clinical operations. Lt. Walders is extending that much more aggressively into the realm of mobile health care and also machine-to-machine types of activities.

For instance, information can be collected from mobile wristbands and fed right back into data collection facilities within the hospital or organization. In some cases they are able to affect clinical appointments in near real time. They are also able to update information in those individuals' records so that they can be advised or counseled appropriately regarding their health care, their physical training regime, appointments and so forth to become more proactive. I think that's an awesome use case and great example that will garner a lot of attention and provide a lot of insight into how big data in particular is capitalizing on some of the new technologies to support the health care environment.

With EMC's help, you can break through your data challenges and use data to drive improved decisions.

DATA.GOV CONTINUES TO EVANGELIZE

"It was an experiment in 2009. We weren't even sure if it was going to be useful. But I think we proved our point."

Jeanne Holm, data evangelist at the General Services Administration, excitedly uttered those words when discussing the progress Data.gov has made in five years of existence. At its outset, the open data hub for government hosted 42 datasets. Now Data.gov is home to nearly half a million sets and sheets of federal, state, local and agency data.

The road to Data.gov's current success, however, has not been without potholes and detours. GovLoop recently sat down with Holm to discuss the challenges of, lessons learned and outlook for opening up our nation's repositories of data.

There are two types of challenges with open data, Holm said. The first is the easier of the two: making data machine- and human-readable and standardized.
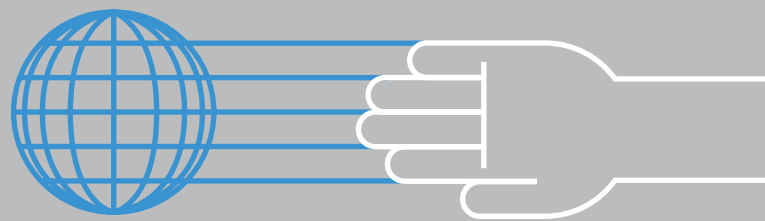
"The more difficult thing that we've made a lot of progress on – but still have a long way to go – is the cultural change within government," Holm said.

Data used to be accessible solely on a need-to-know basis; only departments and agencies that required specific data for research or projects were granted access to government datasets. The public was left in the dark.

> "THE MORE DIFFICULT THING THAT WE'VE MADE A LOT OF PROGRESS ON – but still have a long way to go – IS THE CULTURAL CHANGE WITHIN GOVERNMENT."
>
> *Jeanne Holm, data evangelist at the General Services Administration*

Things are changing, however. "There's been lot of progress from Data.gov's perspective and also from the White House directly with the president's executive orders last year about making open the new default for data," Holm said.

She described two primary benefits to making data open and available that the government and public sector need to realize:

1   People don't have to rely on the government to put data to good use. Open data allows individuals to be more innovative in solving problems. It gives them more access to do remarkable things with data, instead of expecting the government to do it.

2   Open data promotes economic growth for businesses and communities. Several small open data projects have turned into full-fledged businesses. For example, Trulia began as an app to publish community data. Now it's a successful company enabling the public to make smarter purchases, market decisions and investments.

The "Impact" section on Data.gov also tracks economic growth and highlights specific case studies of agencies, organizations and communities making a difference with open data. The site features success stories of companies using open data in the fields of finance, consumption, health, education and energy.

Holm provided her own favorite use cases, too. First she highlighted the work of PatientsLikeMe. Founded by Jamie and Ben Heywood after their brother was diagnosed with amyotrophic lateral sclerosis, and their longtime friend, Jeff Cole, PatientsLikeMe is a website that provides medical data, treatment options and community support to individuals suffering from a variety of conditions. The site connects individuals with shared experiences with conditions such as multiple sclerosis, fibromyalgia, diabetes and Crohn's disease.

Soon after their brother was diagnosed with ALS, the Heywood family began scouring the online medical community for ideas to extend their brother's life. They used open health care data from the Health and Human Services Department and the National Institutes of Health to create a comprehensive online resource for patients and their families.

The online and data-sharing community generates data about the real-world manifestations of diseases, which allows pharmaceutical companies, researchers and medical providers to create more effective products and services.

Holm also recognized the work of M-Farm, a mobile transparency tool for farmers. Throughout Sub-Saharan Africa about 80 percent of people own cell phones, Holm said, which makes SMS messages a prime resource for disseminating information. Farmers can send a message to M-Farm and receive up-to-date market information and agriculture trends in their area.

For instance, farmers can determine which markets are already selling a certain crop in their city and at what prices. Then they can adjust their strategy based on this information, saving time and resources.

This data comes from a variety of sources. M-Farm uses open crop viability and agricultural data from multiple G8 countries. In addition, the release of open datasets from the U.S. Agency for International Development's Famine Early Warning System gave M-Farm access to more than 10 years of historical data about crop prices and market fluctuations.
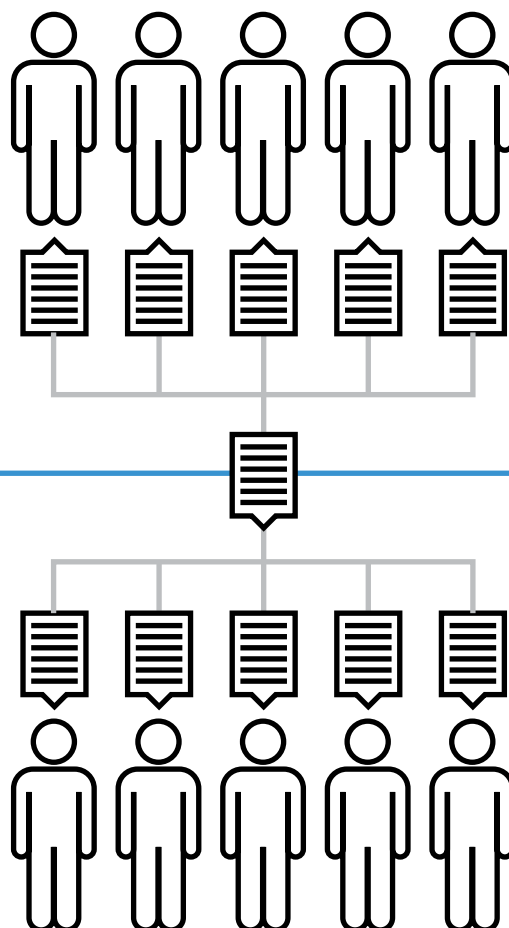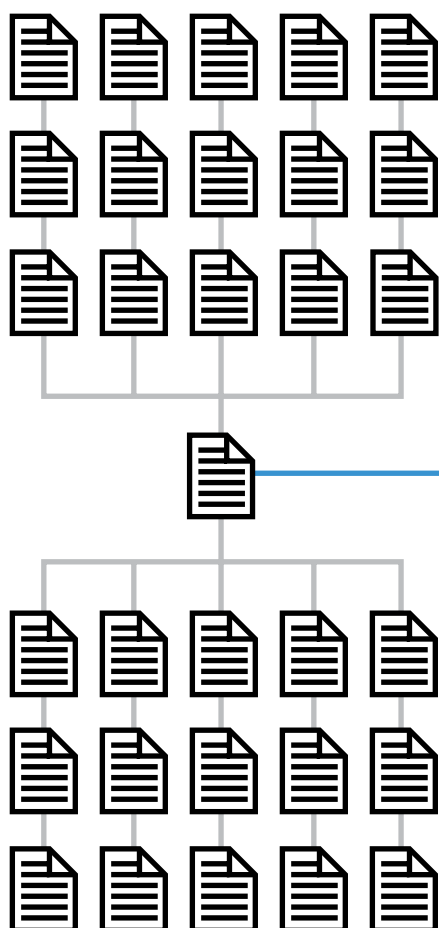
Farmers throughout Kenya are using this innovative web service and open data to increase profits, remain competitive and drive economic growth.

As data evangelist, Holm devotes much of her efforts to educating civil servants and civilians alike about the benefits and potential of open data. Holm believes open data will affect the future of government in four main ways:

1 It will make government more efficient. When data is open and shared, agencies can identify areas of redundancy and streamline datasets. Agencies can combine efforts, saving time and taxpayer money. Also, policymakers can make more informed decisions and work more efficiently with government contractors.
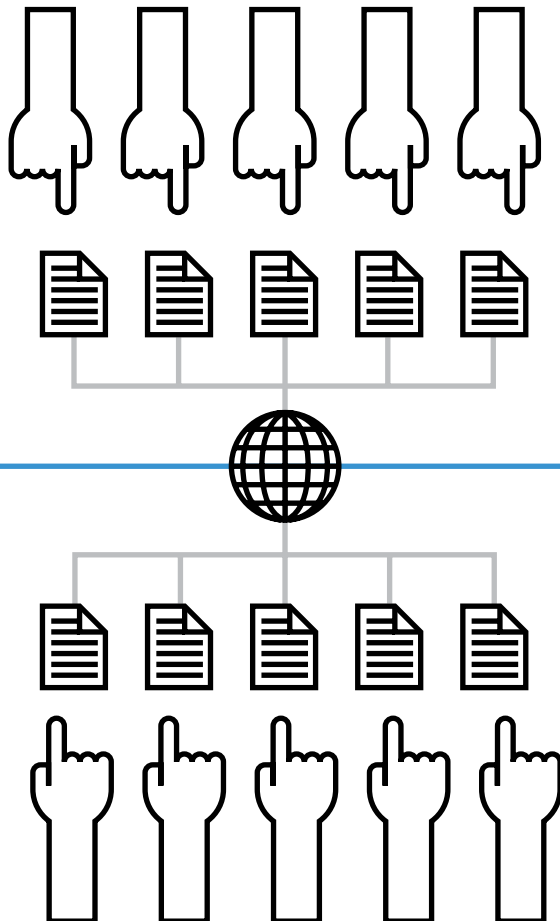
2 It will make government more interactive. Open data allows citizens to become meaningful contributors to government. People can provide government with direct feedback, offer their own insights to datasets and use data in unique ways. Holm related this to the concept of citizen science, where amateurs contribute to research and studies often via crowdsourcing.

3 It will make government more transparent and accountable. One of the main reasons for starting Data.gov was to encourage transparency throughout the public sector, Holm said. In addition, she and her team want the public to be able to identify areas that need to be addressed in government and indicate what data must be open.

4 It will make government more uplifting. "If everyone has access to information democratically," Holm said, "then it gives anyone the power to transform their communities." More people at work with the data paves the way for enhanced opportunities for economic growth, innovation and community development.

Open data can transform how our government operates and how we – as citizens and civil servants – interact with the government. As Holm said, our nation needs to make a commitment to not only be transparent, but live transparently.

YOUR OPEN DATA CHEAT SHEET

Looking to dive into open data? Here's GovLoop's Open Data Cheat Sheet to help you get started.

BEST PRACTICES FROM OUR EXPERTS

Although there are significant challenges to overcome to accelerate adoption of open data, our research found several best practices to help smooth the process. Here are four, based on our survey of 208 public-sector employee and interviews with government and private-sector experts.

IDENTIFY YOUR OPEN DATA CHAMPIONS.

"One of the first things that we did was find the internal advocates and who was already onboard, who were the people that didn't need any cheerleading and were onboard, those were the folks that knew more or less the low-hanging fruit," said Mike Kruger, director of digital engagement at the Department of Commerce. "Once we got our internal champions, we got the low-hanging fruit, some quick wins that helped to guide more strategic questions on our vision."

TALK WITH DATA USERS.

By talking with users of your data, agency officials will know what data they find valuable and how they plan to use it. This will determine the best method of publication, whether the data is released via PDF files, in an application programming interface (API) or through another platform.

"If you open a dataset and no one really knows or can really use it, is it of any real value?" Kruger asked. "It's great to know your stuff is being used. We get to go back and say to the data owners and data producers, 'This is being used for these purposes.' That feels good."

Jeanne Holm recommended being active on social media sites and forums to interact with users. Instead of just broadcasting information, converse with the public about your data and publication methods to gauge feedback and troubleshoot issues. Twitter, StackExchange and GitHub are all valuable social media platforms.

Once a relationship has been established with potential users and data has been published in the ideal formats, it is important to continue to keep users engaged.

GET SENIOR LEADERSHIP BUY-IN.

One of the reasons why Commerce's open data efforts have been so successful is because Secretary Penny Pritzker has been leading the charge. Her private-sector experience has helped her understand how data can transform the agency and helped articulate her vision. Commerce data activists can think about ways to leverage data, knowing they have support from the highest levels.

"Having senior leadership sponsorship is great, and finding that individual is important, because at some point you're going to need money and resources and time commitment to doing things that aren't going to come down at the worker level. That's just not how the system is structured," Kruger said.
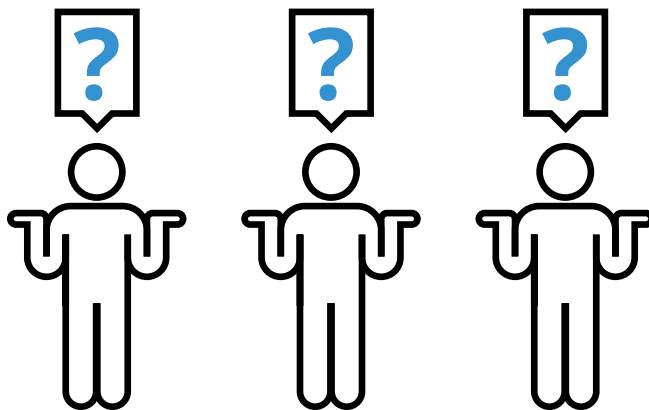
FOCUS ON EDUCATION.

Holm also recommended becoming well versed in open data policies and executive orders before embarking on any open data project. Project Open Data provides comprehensive implementation guidelines for your agency to capture all angles of open data policies. In addition, DigitalGov.gov offers ample training materials.
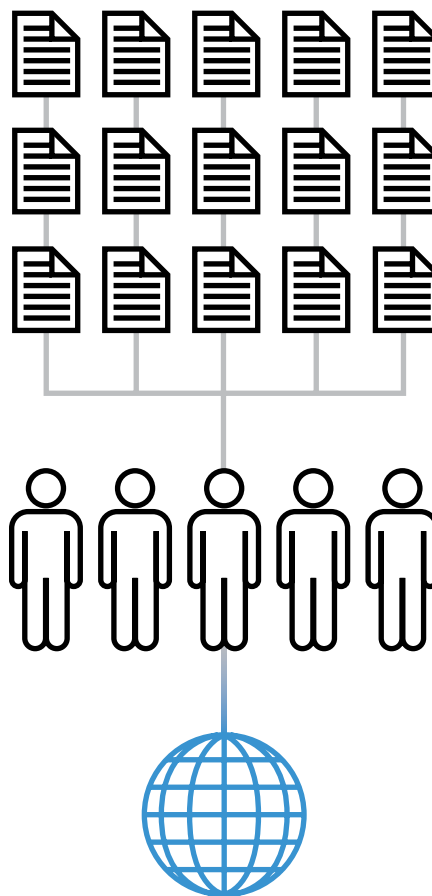
5 QUESTIONS TO ASK BEFORE STARTING YOUR OPEN DATA PROGRAM

1 What is the new value this dataset is providing to our community?

2 How have we engaged with our end users and can we deliver on their needs?

3 Have we conducted an inventory of our data and what is currently public?

4 What are the barriers the developers, citizens or governments face to leverage our data, and how can we fix them?

5 Who are our data champions and how can we gain agency-wide support?

5 BENEFITS OF OPEN DATA

1 ELIMINATES DATA REDUNDANCIES
Agencies can access data from one spot, rather than store and manage it in multiple locations.

2 LEVERAGES DATA OTHER ORGANIZATIONS HOLD
Open data allows individuals, business or other government agencies to connect datasets and find new insights.

3 SUPPORTS DATA-DRIVEN DECISIONS
Open data gives people access to more data, which enables them to make improved decisions on how an agency can be more efficient and effective.

4 SUPPORTS ECONOMIC GROWTH AND FACILITATES NEW SERVICES THAT GOVERNMENT CANNOT INVEST IN
We saw various case studies in this report illustrating how open data allows private-sector companies to create solutions and services that government does not have the resources to create.

5 REDUCES AND CREATES MORE EFFICIENT PROCESSING OF FREEDOM OF INFORMATION ACT REQUESTS
When commonly requested data is freely accessible, fewer FOIA requests have to be submitted and processed.

ABOUT GOVLOOP

GovLoop's mission is to "connect government to improve government." We aim to inspire public sector professionals by serving as the knowledge network for government. GovLoop connects more than 100,000 members, fostering cross-government collaboration, solving common problems and advancing government careers. GovLoop is headquartered in Washington D.C. with a team of dedicated professionals who share a commitment to connect and improve government.

For more information about this report, please reach out at info@govloop.com

GovLoop
1101 15th St NW, Suite 900
Washington, DC 20005

Phone: (202) 407-7421
Fax: (202) 407-7501

www.govloop.com
Twitter: @GovLoop

GLOSSARY & MUST READS

OPEN DATA GLOSSARY

Here's a collection of key terms you need to know to get started with open data. We've pulled a few from Project Open Data. See the complete glossary and source information here.

API: An application programming interface, which is a set of definitions of the ways one piece of computer software communicates with another. It is a method of achieving abstraction, usually (but not necessarily) between higher-level and lower-level software.

GitHub: A social coding platform allowing developers to publicly or privately build code repositories and interact with other developers around these repositories – providing the ability to download or fork a repository, as well as contribute back, resulting in a collaborative environment for software development.

Machine-Readable File: Refers to information or data that is in a format that can be easily processed by a computer without human intervention while ensuring no semantic meaning is lost.

Open Standard: A standard developed or adopted by voluntary consensus standards bodies, both domestic and international. These standards include provisions requiring that owners of relevant intellectual property have agreed to make that intellectual property available on a nondiscriminatory, royalty-free or reasonable royalty basis to all interested parties.

Schema: An XML schema defines the structure of an XML document. It defines things such as which data elements and attributes can appear in a document, how the data elements relate to one another, whether an element is empty or can include text, which types of data are allowed for specific data elements and attributes, and what the default and fixed values are for elements and attributes.

SDK: Software Development Kits are the next step in providing code for developers, after basic code samples. SDKs are more complete code libraries that usually include authentication and production-ready objects, which developers can use after they are more familiar with an API and are ready for integration.

XML: Extensible Markup Language is a flexible language for creating common information formats and sharing both the format and content of data over the Internet and elsewhere. XML is a formatting language recommended by the World Wide Web Consortium.

YOUR OPEN DATA MUST READS

Project Open Data is a great resource to develop open data; you can see more at Project Open Data.

U.S. Government Policy on Open Data – President Obama's memorandum on managing information as an asset, also referred to as "M-13-13."

Implementation Guide – Guidance from the Office of Management and Budget on how to implement M-13-13: "Managing Information as an Asset."

Open Data Principles – Project Open Data defines the principles of what open data consists of.

Standards, Specifications and Formats Supporting Open Data Objectives – This provides best practices for developing data standards at your agency.

Metadata Resources – Some great resources on how to implement common core metadata schemas.

Business Case for Open Data – Here you'll find a list of reasons making the business case for open data, perfect if you are running into cultural roadblocks.

API Basics – This link will teach you the basics of what an API is and how you can leverage one at your agency.

Digital "Personally Identifiable Information" Checklist – A quick resource to help you think about PII, or data that is linked to individuals.

Project Open Data – This is meant to be a living document that offers definitions, case studies, implementation guidance and tools related to open data.

govloop

1101 15th St NW, Suite 900
Washington, DC 20005

Phone: (202) 407-7421
Fax: (202) 407-7501