

Chief Data Officers

Unlocking the Power of Big Data for Federal Agencies

For federal agencies, the veritable explosion of data and high-performance computing power over the last decade represents both promise and potential peril. On the one hand, agencies have unprecedented opportunities to unlock new insights from big data analytics and real-time data streams. On the other, without a robust strategy, many will struggle to adapt and miss valuable opportunities to improve the services they provide citizens and stakeholders.

To address these challenges, many agencies are appointing Chief Data Officers (CDOs). This is a relatively new role and remains largely undefined. Just this February, the White House named DJ Patil as the nation's first Chief Data Officer, adding to a handful of other CDOs serving at the Departments of Commerce and Transportation, the FCC, and FEMA, among others.¹ With the tremendous volume, velocity, and variety of data available to today's policymakers, the addition of a CDO can assist agencies in developing strategies to merge silos, standardize data, and manage it for effective use toward promoting the organization's goals.

Treating Data as an Asset

According to former Defense Information Systems Agency (DISA) enterprise architect and data science expert, Peter Aiken, the objective for federal agencies should be "to start treating data as an asset rather than as a by-product."² Treating data as an asset means taking ownership of the data and understanding that it has inherent, quantifiable value — to your organization and others.³ Too often, says Aiken, a one-off approach to data management relying on varying data standards leaves small piles of data scattered across business units and agencies that are fundamentally useless

to one another. By integrating these data silos under a master data management strategy, agencies can unlock enormous business value.

However, the prospect of devising and implementing an aggressive cross-agency data strategy presents a major challenge for most public sector chief information officers (CIOs), especially given the heavy burden already expected of them. According to Micheline Casey, Chief Data Officer at the Federal Reserve Board, technological shifts over the last decade have forced CIOs to function as an "infrastructure officer" more so than an "information officer," a trend that has opened the door for the CDO. "You need someone who's accountable for defining the business rules, processes, and best practices around particular data sets and thinking about the data strategy vis-a-vis the mission of the organization," she says.⁴

Unlocking the Value of Your Agency's Data

Unlocking business value from data hinges on an agency's ability to craft a viable data management strategy and to match that strategy with the buy-in and resources necessary to implement it successfully. Leadership — whether provided by a CDO, CIO, or other executive roles — can be essential in this regard. But,

"Start treating data as an asset rather than as a by-product."

Peter Aiken

Author, *The Case for the Chief Data Officer*

what goes into an effective data management strategy? A 2014 study by the IBM Institute for Business Value provides a useful framework. According to the study, the most important aspects of a data management strategy are *data leverage*, *data exchange*, *data enrichment*, *data upkeep*, and *data protection*.⁵ Most federal agencies are already taking the majority of these steps; however, the real challenge is coherently orchestrating them so that each process reinforces the others.

The Five Keys to an Effective Data Management Strategy

One of the primary objectives of data management is ensuring that each agency is able to leverage the entire universe of data available to it in order to drive decision making. But for most federal agencies, says Anthony Scriffignano, Chief Data Scientist at Dun & Bradstreet, the problem isn't a lack of data – it's the inability to organize existing data in a way that enables officials to derive value from the information available. "We're drowning in data," says Scriffignano, "Our problem is not, 'How do we get more data?' Our problem is, 'How do we make sense of it? How do we get value from it?'"⁶

CDOs can lead efforts to merge silos, disseminate best practices, and set business rules for their organizations,

The 5 Principles of Data Management

Data leverage – Finding ways to use existing data assets to improve decision making and advance the cause of the organization

Data exchange – Finding ways to share and exchange data assets across agencies to advance mission objectives and to open data to the public

Data enrichment – Supplementing existing data with fragmented internal sources or integrating it with innovative or non-traditional data streams

Data upkeep – Managing the health of the organization's data and maintaining its integrity, value, and timeliness

Data protection – Safeguarding the organization's proprietary data against unauthorized access and dissemination

but when it comes down to it, one of their most important functions will be to maximize data quality by reducing data that is Redundant, Obsolete, and Trivial (ROT).⁷ Take U.S. Census data on small businesses, for instance. Over time, businesses may open and close,

"We're drowning in data. Our problem is not, 'How do we get more data?' Our problem is, 'How do we make sense of it?'"

**Anthony Scriffignano, Chief Data Scientist,
Dun & Bradstreet**

change addresses, expand and contract. Unless there are protocols in place to continuously update that data, the accuracy of analyses derived from that data may suffer. Further, by removing redundant or trivial data sources, agencies can drive down storage costs while ensuring data quality.

In recent years, federal agencies are being asked to go a step beyond leveraging their own data effectively. In a 2013 Executive Order, President Obama mandated that all agencies promote data interoperability and make government data sources accessible and usable to other government entities and the public.⁸ The ability to share data effectively could produce tremendous positive impacts for transparency and economic opportunity. Nevertheless, the requirement that all agencies make their data machine-readable by default introduces its own set of challenges, including the need to adopt new formats such as RDF, XML, and JSON as well as develop APIs and keep extensive metadata logs.⁹

Improving data interoperability could also help agencies achieve their core objectives in more cost-effective ways – for instance, through the facilitation of public-private partnerships. For Scott Shoup, Chief Data Officer at FEMA, one major priority is structuring his agency's data to streamline collaboration with the private companies, humanitarian groups, and non-profit organizations that act as force multipliers for FEMA in disaster zones. "We have to be able to do our data collection and push that data out to our partners to the largest extent possible so that FEMA isn't everything to everyone," says Shoup, "Instead of bringing trucks of water to a disaster, we help big-box stores clear the roads so that they bring the water."¹⁰

In addition, the data universe available to federal agencies is expanding at an exponential rate. Yet, many continue to struggle with how to tap into unstructured or real-time data streams and use them to enrich the efficiency and effectiveness of their legacy data sources. “We’re dealing with a lot more unstructured data than the [Federal Reserve] Board has ever been used to,” notes Casey, whose team is currently using text analytics to pull keywords out of unstructured data sources, like financial documents, and pairing them with raw, structured stress-test data on financial institutions.¹¹

The Federal Reserve is also working to collect real-time data from nationwide car-sharing services, like Uber, and analyzing it alongside conventional economic indicators to gain a more granular picture of the automobile market. “When analyzing high-frequency and non-traditional data streams, the symbol I like to use is, we’re trying to get a ‘fit-band for the economy,’” she explains, “Just like what a fit-band tells you about what’s going on with your body, we’re trying to understand the health and well-being of the economy in near real-time.”¹²

Although the primary responsibility of the CDO is to promote data quality and interoperability, recent

data breaches involving the personal information of millions of government employees illustrate that federal agencies must also be held accountable for securing their sensitive and proprietary data against unauthorized access or disclosure. In this regard, CDOs can play an important role in coordinating with Chief Information Security Officers (CISOs) to implement data encryption, strict identity and access management, and multi-factor authentication.¹³

Asking the Right Questions

Today’s federal CDO must be more than simply an agency’s “chief data scientist.” He or she must embrace the role of “chief data evangelist” within the organization and possess a range of skills that include business acumen, interpersonal communication, technological understanding, and a keen eye for process. More importantly, federal agencies need future CDOs capable of outside-the-box thinking and who understand that asking the right question is oftentimes just as valuable as finding the right answer. “There’s no substitute for a passionate question,” remarks Department of Transportation Chief Data Officer Dan Morgan, “because that will organize what we do around data.”¹⁴

Sources

1. Jack Moore, “Rise of the Data Chiefs: Meet the Federal Officials Aiming to Usher in Government’s ‘Golden Age’ of Data.” *Nextgov*: March 18, 2015
2. Peter Aiken, “Rise of the Chief Data Officer.” A live event presented by *Nextgov* and Dun & Bradstreet: June 23, 2015
3. Nicole Laskowski, “Infonomics Treats Data as a Business Asset.” *Tech Target*:
4. Micheline Casey, “Rise of the Chief Data Officer.” A live event presented by *Nextgov* and Dun & Bradstreet: June 23, 2015
5. Teerlink, Sigmon, Gow, and Banerjee, “The New Hero of Big Data and Analytics: The Chief Data Officer.” *IBM Institute for Business Value*: June 2014; The IBM report uses the term data monetization rather than exchange, given its intended consumption by a private sector audience.
6. Anthony Scriffignano, “Rise of the Chief Data Officer.” A live event presented by *Nextgov* and Dun & Bradstreet: June 23, 2015
7. Aiken, “Rise of the Chief Data Officer.” 6/23/2015
8. President Barack Obama, “Executive Order 13642: Making Open and Machine-Readable the New Default for Government Information.” May 9, 2013
9. Jim Hendler, “A Primer on Machine Readability for Online Documents and Data.” *Data.gov*: September 24, 2012
10. Scott Shoup, “Rise of the Chief Data Officer.” A live event presented by *Nextgov* and Dun & Bradstreet: June 23, 2015
11. Casey, “Rise of the Chief Data Officer.” 6/23/2015
12. Casey, “Rise of the Chief Data Officer.” 6/23/2015
13. Hallie Golden, “Not Just OPM — Agency Cybersecurity Incidents on the Rise.” *Nextgov*: July 10, 2015
14. Dan Morgan, “Rise of the Chief Data Officer.” A live event presented by *Nextgov* and Dun & Bradstreet: June 23, 2015

About Dun & Bradstreet

Dun & Bradstreet (NYSE: DNB) grows the most valuable relationships in business. By uncovering truth and meaning from data, we connect customers with the prospects, suppliers, clients and partners that matter most, and have since 1841. Nearly ninety percent of the Fortune 500, and companies of every size around the world, rely on our data, insights and analytics. For more about Dun & Bradstreet, visit dnb.com/government.

About GBC

Government Business Council (GBC), the research arm of Government Executive Media Group, is dedicated to advancing the business of government through analysis and insight. GBC partners with industry to share best practices with top government decision makers, understanding the deep value inherent in industry’s experience engaging and supporting federal agencies.