Government Business Council

dun & bradstreet

From Clay to Bricks

Shaping Federal Data into an Asset

"Data! Data! Data!" cries Sherlock Holmes during a case, "I can't make bricks without clay."[1] The erudite detective would have been pleased with the astonishing and unprecedented surge in data production that characterizes our current age. By 2020, the amount of data generated annually is projected to reach 44 zettabytes — a stunning figure considering that, until 2013, humanity had produced a total of 2.7 zettabytes.[2] These days, federal leaders have a world of clay at their fingertips – the question, then, is how to best mold that clay into bricks. Possessing a surplus of data without also developing a plan of action is more of a liability than an asset, and the key challenge ahead for federal agencies is how to successfully glean valuable insights from the data. Data is central toward informed decision making, and in order to drive mission objectives, organizations must critically evaluate their data management and utilization framework in order to leverage information as a key strategic asset.

Data in Today's Government

Public and private sector decision-making rests on the availability and accuracy of data, and the federal government is increasingly recognizing data's immense potential in driving critical insights and fulfilling mission objectives. The 2013 Office of Management and Budget (OMB) memo refers to data as a "valuable natural resource" and calls on departments and agencies to "manage information as an asset throughout its life cycle to promote openness and interoperability, and properly safeguard systems and information."[3]

In their effort to further prioritize data, organizations are launching both external and internal initiatives: the Environmental Protection Agency (EPA) is streamlining its data collection processes through an application programming interface as well as transitioning to cloud-based data storage and access;[4] the Office of Science and Technology Policy's (OSTP) newly-established "Data Cabinet" brings together federal data leaders each month to share challenges, ideas, and best practices;[5] and the Federal Communications Commission (FCC) is pushing for greater integration between the "custodial" aspects of data management (e.g., security and infrastructure-maintenance) and the "fun," innovative side of data. As FCC Chief Information Officer (CIO) David Bray puts it, "How can we bring [different IT camps] together so that there's a shared sense of fun and a shared sense of responsibility?"[6]

The federal government is also increasingly cognizant of the potential for open data to help drive continuous innovation — as former U.S. Deputy Chief Technology Officer (CTO) Beth Noveck notes, making data public and participatory catalyzes the transformation of a passive, "read-only society" into "a writable society where we have the power to change our communities, to change our institutions."[7] In keeping with President Obama's Open Government Initiative,[8] Data.gov now features over 200,000 datasets on education, public safety, healthcare, energy, agriculture, and more — an open, exploratory ecosystem where citizens can channel their values, vision, and expertise into creating a government that is truly "by the people, for the people."[9]

Reframing Data for the Future

However, even with expanded resources and growing awareness of the importance of data, federal agencies often struggle to determine how to most effectively employ data as a problem-solving asset. According to General Services Administration (GSA) Deputy Administrator Adam Neufeld, the key is to be

deliberate about both starting and abandoning data efforts. In order to do this, organizations need to reframe data as being part of an entire cycle — and to determine whether they have an actionable vision for each step of this cycle before embarking on any data initiative.[10]

The first step of this cycle is to clearly establish the problem at hand before brainstorming possible solutions. While this may seem self-evident, knowing exactly what to look for is essential in order to isolate valuable insights from excess data — and to help evaluate whether the problem requires data in the first place. Second, agencies need to establish the best data collection process and assess whether more data is really needed. According to David Bray, the problem often isn't lack of data, but siloed data. Thus, the next step for the federal government is creating a common platform that will allow different organizations to build from and mix shared data — as Bray puts it, "Just because you're the Department of Agriculture doesn't mean you aren't also intersecting with the EPA or HHS or DoD."[11]

Once organizations collect the data, they need to address storage. Ensuring that employees receive adequate data management training is an essential part of this process, but according to a September 2015 Government Business Council flash poll, many agencies have significant room for improvement on this front: only 45% of federal employees agree or strongly agree that their organization provides adequate records management training.[12] However, effective, comprehensive storage processes are critical for both security and innovation — and while organizations recognize the importance of opening data, they are also, as EPA Chief Data Scientist Robin Thottungal notes, profoundly risk-averse.[13] By initially setting parameters on and then building a governance structure around access, organizations can simultaneously mitigate risk and enable openness: the GSA's Data to Decisions (D2D) storage platform, for instance, is a cloud-based, open-source platform designed to both maintain security and make data available to stakeholders.[14]

Once collected and stored, data needs to be analyzed — a step that requires skill and methodological integrity. "Statistics are no substitute for judgment," cautioned statesman Henry Clay, and in order to draw thorough, honest insights, organizations need to ensure that they possess a talented data workforce with the requisite skills as well as data-literate leaders.[15] From there, decision makers assess the data and analyses in order to make an informed decision. At this point, says Dun & Bradstreet Chief Data Scientist Anthony Scriffignano, organizations need to take what he calls the "dispositive threshold" into account: factoring both available and unavailable data into the decision. It's possible to have enough data to make a decision, but not necessarily enough to make a *good* decision: leaders should bear this in mind, recognize and consider the degrees of freedom in any analysis, and factor in the probable impact of inaccessible data.[16] But above all, agencies need to get creative, envision the possibilities in data, and "use their business of data most effectively and most efficiently to help the public."[17]

Finally, all this analysis and vision only means something once it's implemented. In order to bring an idea to fruition, organization leaders need to successfully convey the rationale and potential benefits of an idea to stakeholders. The GSA, for example, collected transaction data from an office supply contracting vehicle, and after analyzing prices across the board, came to the conclusion that vendors didn't have a sense of the prices their competitors were offering. They then created an automated tool that notified vendors when their offered price was significantly above average and thus uncompetitive — something that, once communicated to contractors, received enthusiastic buy-in and participation. The GSA's initiative has yielded 1,397,457 voluntary price reductions to date, generating average savings of 27.1% — an impressive example of data in action.[18]

The Road Ahead

"Now, here, you see, it takes all the running you can do, to keep in the same place," the Red Queen tells Alice in the looking-glass world. "If you want to get somewhere else, you must run at least twice as fast as that!"[19] According to Anthony Scriffignano, the federal government is caught up in its own iteration of this Red Queen problem: besieged with growing demands, many agencies are falling into the trap of devoting increasing amounts of data to problems without first devising a holistic strategy — in effect, running faster and faster without getting anywhere.[20] But data-driven government isn't simply defined by the presence of data; it's propelled by a fundamental shift in thinking. By thinking holistically about how data can be shaped into an asset, organizations can more confidently and efficiently leverage this powerful resource to address the emerging challenges of the twenty-first century.

Sources

1 Arthur Conan Doyle, *The Adventures of Sherlock Holmes*. 1892.

2 *Information Age*, "4 key Internet of Things analytics trends for marketing pros." June 9, 2015. http://www.information-age.com/it-management/strategy-and-innovation/123459628/4-key-internet-things-analytics-trends-marketing-pros

3 Executive Office of the President, "Memorandum for the Heads of Executive Departments and Agencies." May 9, 2013. https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf

4 GCN.com, "EPA looks to improve data aggregation with APIs." May 20, 2016. https://gcn.com/articles/2016/05/20/epa-data-aggregation.aspx

5 Natalie Harris, *Nextgov* event: "Data in Action: Improving the Business of Government." June 23, 2016. http://www.nextgov.com/feature/data-action-livestream/

6 David Bray, *Nextgov* event: "Data in Action: Improving the Business of Government."

7 Beth Noveck, "Demand a more open-source government." June 2012. https://www.ted.com/talks/beth_noveck_demand_a_more_open_source_government

8 Whitehouse.gov, "Open Government Initiative." https://www.whitehouse.gov/open

9 Office of Management and Budget, "Analytical Perspectives, Budget of the United States Government, Fiscal Year 2017: Information Technology." https://www.whitehouse.gov/sites/default/files/omb/budget/fy2017/assets/spec.pdf

10 Adam Neufeld, *Nextgov* event: "Data in Action: Improving the Business of Government."

11 David Bray, *Nextgov* event: "Data in Action: Improving the Business of Government."

12 Government Business Council, "Flash Poll Series: Information Management." September 16, 2015. http://www.govexec.com/gbc/flash-poll-center-information-management/?oref=insights-river

13 Robin Thottungal, *Nextgov* event: "Data in Action: Improving the Business of Government."

14 Adam Neufeld, *Nextgov* event: "Data in Action: Improving the Business of Government."

15 Adam Neufeld, *Nextgov* event: "Data in Action: Improving the Business of Government."

16 Anthony Scriffignano, *Nextgov* event: "Data in Action: Improving the Business of Government."

17 Natalie Harris, *Nextgov* event: "Data in Action: Improving the Business of Government."

18 Adam Neufeld, *Nextgov* event: "Data in Action: Improving the Business of Government."

19 Lewis Carroll, *Through the Looking-Glass, and What Alice Alice Found There*. 1871.

20 Anthony Scriffignano, *Nextgov* event: "Data in Action: Improving the Business of Government."

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.

About GBC

As Government Executive Media Group's research division, Government Business Council (GBC) is dedicated to advancing the business of government through analysis, insight, and analytical independence. An extension of Government Executive's 40 years of exemplary editorial standards and commitment to the highest ethical values, GBC studies influential decision makers from across government to produce intelligence-based research and analysis.