The Big Data Gap

May 7, 2012

Underwritten by:

NetApp

www.meritalk.com
At the end of the day, agency and IT leaders want technology to support overall government mission success, not just IT outcomes. Better use of extensive data assets – harnessing big data – is at the core of how IT can improve the execution of government. Are agencies ready?

To better understand how agencies are positioned to use data – specifically big data – to support their agencies, MeriTalk surveyed 151 Federal IT professionals to determine if they have the necessary technology, management, and talent.

The study highlights the big data opportunity and outlines the gap between big data possibility and big data reality for most agencies.
Agency executives want IT to enable fundamental mission changes. Unlocking and harnessing agency data can deliver that progress:

- **Top benefits** include: Improving overall efficiency (59%); speed/accuracy of decisions (51%); and the ability to forecast (30%)

Yet most of that promising progress is locked away in unused or inaccessible data:

- **Just 60%** of IT professionals say their agency is analyzing the data they collect and less than half (40%) are using their data to make strategic decisions
- IT professionals report, on average, that it will take their agency at least **3 years** to fully take advantage of big data

Whether it is an opportunity or a challenge, data continues to **grow**:

- **87%** of IT professionals say their stored data has grown in the last two years
- **96%** expect their data to grow in the next two years (by an average of 64%)

The big data gap:

- Ownership of data is not clear. **42%** believe that IT owns the data; **12%** believe ownership belongs to the C-level
- **31%** of data is unstructured, and that amount is increasing
- Agencies estimate they have just **49%** of the data storage/access; **46%** of the computational power; and **44%** of the personnel they need to leverage big data and drive mission results
Big Data = Better Government

• IT professionals believe they can use big data to make their agency more efficient and enable leadership to make better decisions.

In your opinion, what are the top three advantages to successfully managing big data?*

- Improving overall agency efficiency: 59%
- Improving speed/accuracy of decisions: 51%
- Ability to forecast: 30%
- Ease of identifying opportunities for savings: 25%
- Greater understanding of citizen needs: 23%
- Greater understanding of our agency: 21%
- Increasing transparency: 21%

*Respondents asked to select up to three responses

Take Away: Big Data, Big Benefits
While agencies agree there are many benefits to big data, the technology and its applications are still emerging.

**Where is your agency with big data?**

- **Civilian agencies**
  - **We are not discussing big data**: 9%
  - **We are learning about it and how it could work for us**: 60%
  - **We are designing a plan and/or proof of concept**: 17%
  - **We have a plan/test lab to implement new processes**: 8%
  - **We have infrastructure/processes to successfully leverage it**: 6%

- **DoD/Intel agencies**
  - **We are not discussing big data**: 31%
  - **We are learning about it and how it could work for us**: 42%
  - **We are designing a plan and/or proof of concept**: 19%
  - **We have a plan/test lab to implement new processes**: 5%
  - **We have infrastructure/processes to successfully leverage it**: 3%

Agencies report, on average, it will take them at least **3 years** to take full advantage of big data.

**Take Away:** Work to be Done. Gains over the Horizon.
In fact, agencies are doing very little with their data.

**What is your agency currently doing with the large amounts of data it collects?***

- Capturing the data: 60%
- Analyzing the data: 60%
- Measuring the volume of data: 43%
- Making strategic decisions with the data: 40%
- None of the above: 3%

**How does your agency measure or analyze this data?**

- Execute historical analysis of data: 60%
- Categorize data: 49%
- Seek out trends in the data sets: 49%
- Seek answers to a specific problem by searching through the data: 46%
- Execute real-time data analysis: 42%
- Collaborate with other agencies to analyze shared data: 28%

*Respondents asked to select all that apply  **Respondents who said they were capturing, analyzing, measuring, or making strategic decisions with data asked to select all that apply.*

**Take Away:** Agencies Spend to Collect but Are Yet to Unlock the Return
Agencies currently store an average of 1.61 petabytes of data, but expect to get to 2.63 petabytes in just the next two years.

### In the last two years...

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>87%</td>
<td>61%</td>
</tr>
<tr>
<td></td>
<td>of agencies say their stored data has grown</td>
</tr>
</tbody>
</table>

### In the next two years...

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>96%</td>
<td>64%</td>
</tr>
<tr>
<td></td>
<td>of agencies expect their stored data to grow</td>
</tr>
</tbody>
</table>

One petabyte of data is the equivalent of **20 million** four-drawer filing cabinets filled with text.

---

**Take Away:** Agencies Need More – and More-Efficient – Data Storage

---

*According to respondents who indicated their data has grown over the last two years, n=131 **http://www.whatsabyte.com/
Agencies will require new approaches and new solutions to combine structured and unstructured data.

Has the amount of unstructured data you store increased over the past two years?

- **Yes**: 31%
- **No**: 64%
- **Unsure**: 5%

On average, agencies report 31% of their data is unstructured.

**Take Away:** Agencies Can No Longer Rely on Simple, Structured Data
Management Hurdles

• Forget the technology, what about the basic management issues? Agencies can’t even agree on who owns the data.

Who owns the data?*

<table>
<thead>
<tr>
<th>Category</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT</td>
<td>42%</td>
</tr>
<tr>
<td>Department that generates data</td>
<td>28%</td>
</tr>
<tr>
<td>C-Level</td>
<td>12%</td>
</tr>
<tr>
<td>Data analysts</td>
<td>10%</td>
</tr>
<tr>
<td>Data scientists</td>
<td>5%</td>
</tr>
<tr>
<td>Other</td>
<td>3%</td>
</tr>
</tbody>
</table>

“One of the biggest challenges to successful big data management is the coordination of multiple disparate ends when developing and leveraging data analysis.” – DoD CIO/CTO

Take Away: Solve the Management Issues First

*Respondents asked to select one
What are the most significant challenges your agency faces when it comes to managing large amounts of data?

- Storage capacity: 40%
- Distribution & sharing: 36%
- Search/retrieval: 35%
- Speed of analysis/processing (takes too long): 34%
- Analyzing: 32%
- Capture: 19%

**Take Away:** Missing Key Parts of the Puzzle

*Average percentages based on respondent feedback **Respondents asked to select top three challenges*
Unmanageable Data

• More than half of agencies report a break in their system, saying at least one dataset has outgrown their current management tools.

57% of agencies say they have at least one dataset that’s grown too big to work with using their current data management tools and/or infrastructure*

*Of those familiar with their dataset sizes

Take Away: Agencies Either Need More Computing Storage or to Use What They Have More Efficiently
While the majority say that their agency can scale to meet data growth, they estimate it would take them 10 months to do so.

64% of IT professionals say their agency’s data management system can be easily expanded/upgraded on demand.

However… they estimate 10 months as the average time they could double their short- to medium-term capacity.

Take Away: Opportunity for Cloud?
• Some agencies are taking steps to optimize their data today and imagine a better, faster, less wasteful future

What steps is your agency taking to improve its ability to manage and make decisions with big data?*

1. Invest in IT infrastructure to optimize data storage (39%)
2. Train IT professionals to manage/analyze big data (33%)
3. Improve the security of stored data (31%)
4. Educate senior management on big data issues (28%)
5. Solicit contractor support for big data initiatives (25%)
6. Invest in IT systems/solutions to optimize data processing (25%)

How could big data help your agency better perform its mission?

“It can determine areas to target for the future; staffing; budgetary issues.” – Federal Civilian IT Director/Supervisor

“It will help us become more transparent and most of all save money for taxpayers in a receding economy.” – Federal Civilian Data Center Manager

“If it gave us the ability to make better projections, then our simulations would be more accurate.” – DoD IT Manager

“It can detect and/or predict problems before they occur.” – Federal Civilian CIO/CTO

Take Away: Bright Future if Agencies Take the Right Steps

*Respondents asked to select all that apply
Recommendations

• Tackle the management issue first: determine which department owns big data

• Start preparing for new solutions and approaches to unstructured data

• Look down the road and make decisions today that work with using big data tomorrow

• Expand your data management plan; more data is coming
• MeriTalk, on behalf of NetApp, conducted an online survey of **151 Federal government CIOs and IT managers** in March 2012. The report has a margin of error of ± 7.95% at a 95% confidence level.

**IT manager titles:**

<table>
<thead>
<tr>
<th>%</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>6%</td>
<td>Chief Information Officer/Chief Technology Officer</td>
</tr>
<tr>
<td>4%</td>
<td>Deputy Chief Information Officer/Deputy Chief Technology Officer</td>
</tr>
<tr>
<td>22%</td>
<td>IT Director/Supervisor</td>
</tr>
<tr>
<td>40%</td>
<td>IT Manager</td>
</tr>
<tr>
<td>2%</td>
<td>Network Administrator</td>
</tr>
<tr>
<td>5%</td>
<td>Data Center Manager</td>
</tr>
<tr>
<td>3%</td>
<td>Data Scientist</td>
</tr>
<tr>
<td>2%</td>
<td>Research/Lab management</td>
</tr>
<tr>
<td>16%</td>
<td>Other IT management</td>
</tr>
</tbody>
</table>

**Agency type:**

<table>
<thead>
<tr>
<th>%</th>
<th>Agency Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>Federal Civilian</td>
</tr>
<tr>
<td>50%</td>
<td>Department of Defense or Intel</td>
</tr>
</tbody>
</table>

**100% of respondents** are familiar with their agency’s data capture, management, and analysis capabilities.
Thank You

Erin Leahy – MeriTalk
eleahy@meritalk.com
(703) 883-9000 ext. 139

www.meritalk.com