Big Data: A Big Deal for Public Sector Organizations

What's the big deal about big data? Whether they're federal, national, defense, state or local, organizations around the world see the ability to analyze large volumes of diverse data as an important new tool for reducing costs, improving efficiency, and delivering more effective services to constituents.

The promise of better financial and performance outcomes thanks to successful big data strategies helps explain why even during times of constrained budgets, governments around the world are eyeing investments in big data. The Obama Administration announced a Big Data Initiative\(^1\) on March 29, 2012 with over $200 million in new spending this year to improve the tools needed to process and visualize huge volumes of digital data.

Governments around the world see big data as an important new tool to help them meet their mission goals during tough economic times

However, managing and capitalizing on big data can be challenging. The keys are to first identify a big data use case that directly impacts the agencies’ mission and second, ensure that the scope is not too broad. Big Data architectures are designed to scale, so starting small and growing is easier than with other approaches. Next, project managers need to identify integrated, end-to-end solutions that fully address big data's distinctive characteristics. With that information in hand, public-sector IT managers can develop clear financial justifications for big data solutions by demonstrating their value to the overall mission, to constituent services, and to the internal processes that power the organization. To achieve these goals, IT administrators should evaluate commercial big data options according to their ability to capture and organize a wide variety of data types from different sources. The best solutions will make it easy for organization personnel to integrate into their existing workflow and to simultaneously analyze large volumes of diverse information within the context of all the traditional data the organization currently manages.

Big Payoffs

Fortunately, the due diligence required to develop an effective big data strategy has the potential to deliver significant payoff for the organization.

---

\(^1\) [http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf)
Research by the McKinsey Global Institute\(^2\) shows that big data solutions could potentially reduce administrative costs in Europe's public sector by up to 20 percent, creating the equivalent of €300 billion (US$446 billion) in new value. The savings would come from both efficiency gains and more effective collection activities for tax revenues.

Similarly, if the U.S. healthcare sector, including publicly supported operations, used big data to improve efficiency and care quality, the sector could create more than US$300 billion in value year after year. Two-thirds of that would be from potential reduced healthcare expenditures of about 8 percent, according to the McKinsey Global Institute.

But the benefits of big data to the public sector aren't only about finances. The right strategy can also help agencies more effectively serve their constituents. For example, Isidore Sobkowski, chief information officer for New York City's Department of Health & Human Services, says big data offers a way to improve citizen services by creating a 360-degree view of each constituent, and providing that information to the appropriate government caseworkers\(^3\). Social services organizations are an example of a public sector function that can use large and diverse data sets to move beyond standard reports that address basic questions about a constituent. With predictive analytics case managers can ask, “What can I do for this person, at this moment, to achieve the best long term outcome?”

**What is Big Data?**

Because the big data phenomenon is still relatively new, it’s important for managers to understand what makes this trend unique. Consider the four “Vs” that define the key characteristics of big data:

- **Volume** – Typically, big data refers to high volumes of information, such as the petabytes of data that is generated by geospatial applications used by state and local governments for mapping and situational-awareness activities. But “big” often is in the eye of the beholder. Big data solutions can also help organizations improve citizen services by collecting and analyzing the relatively small amounts of information, such as those generated in direct social-media conversations with an organization.

- **Velocity** – Like water from a fire hose, data is constantly flowing into agencies at intense rates. Without solutions designed to manage the accelerating rate of data creation and use, in these volumes agencies risk not becoming aware of information that could be critical to their missions or internal operations.

- **Variety** – OLTP (Online Transaction Processing) data is relatively well described and has a fixed schema. By contrast, non-traditional data sets associated with big data, such as tweets, blogs, wiki entries, videos, and audio files, are becoming common data formats in many types of internal and constituent-facing applications used by organizations today.

- **Value** – High volumes of data produced at accelerated rates typically mean organizations contend with mass quantities of extraneous or low-value information. This makes finding important, and perhaps life-saving, insights difficult. Big data addresses this challenge and gives public safety, law enforcement, and other types of agencies new tools and techniques to extract knowledge from information.


\(^3\) [http://issuu.com/govloop/docs/oracle_industry_perspective_single/1?mode=window](http://issuu.com/govloop/docs/oracle_industry_perspective_single/1?mode=window)
Clear Benefits

Organizations that find success addressing these four characteristics of big data have the potential to deliver new services, reduce costs, optimize existing IT investments, and even save lives. Potential use cases include:

- **Constituent sentiment** – Tracking information on websites, blogs, twitter feeds and mainstream media sources can help policy makers prioritize new services initiatives or uncover potential areas of civil unrest.

- **Fraud detection and prevention** – Revenue collectors and social-service organizations can more effectively track and analyze citizen activities to spot abnormal behavioral patterns that may indicate costly mistakes or fraud.

- **Economic analysis** – By correlating multiple sources of data, government economists can better weather today’s volatility with more accurate financial forecast.

- **Healthcare** – Collecting and analyzing public health trends and patient records can help organizations proactively respond to health threats while also improving patient care, and reducing fraud.

- **Open Government** – The free-flow of information from organizations to citizens promotes greater trust between citizens and government.

- **Tax collections** – Organizations can integrate structured and unstructured data from social media and other sources to validate information or flag potential audits.

- **Threat identification** – Searches of financial, news, and social media (wikis, blogs, websites, Twitter) content, can help organizations uncover correlations in funding and transportation of hazardous materials.

- **Cyber Security** – Big Data solutions can collect, organize and analyze the vast amounts of data from government computer networks to give cyber defenders greater ability to detect and counter malicious attacks.

**Oracle’s Big Data Edge**

Oracle is uniquely qualified to deliver all the technology components needed to address the big data challenge – including software and hardware – as one complete, integrated system. Of particular interest to budget-constrained public-sector organizations, Oracle’s big data strategy enables them to build on their current investments in organization-wide data architectures rather than forcing them to start from scratch as they roll out big data solutions.
FIGURE 1: A Complete, Engineered Solution

The Oracle Big Data Appliance, in conjunction with Oracle Exadata Database Machine and the Oracle Exalytics Business Intelligence Machine, delivers everything organizations need to acquire, organize, analyze and maximize the value of big data.

Oracle’s enterprise-class, end-to-end big data Platform provides a complete data architecture across the four essential steps of a successful big data strategy: acquire, organize, analyze, and decide.

**Step #1: Acquire.** The Oracle Big Data Appliance combines leading open source technologies with software developed by Oracle to meet big data requirements. The Oracle Big Data Appliance comes in a full rack configuration with 18 Sun servers from Oracle that offers a total storage capacity of 648 terabytes. The Oracle Big Data Appliance also includes both the Oracle NoSQL Database and the Hadoop Distributed File System (HDFS) to provide mechanisms to acquire and organize massive volumes of unstructured data in the context of enterprise architecture. Oracle’s engineered systems eliminate many configuration and integration steps to speed deployments.

**Step #2: Organize.** Once the Oracle Big Data Appliance acquires the data, it can then process the information using Hadoop’s MapReduce capabilities. The output can then be loaded into the Oracle Database via the Oracle Loader for Hadoop or it can be accessed from the Database via the Oracle Direct Connector for HDFS. Once the results are in the Oracle Database, the information can be merged and enhanced by other structured data that may have come from existing traditional business systems.

**Step #3: Analyze.** The Oracle Exadata Database Machine provides outstanding performance for hosting data warehouses and for OLTP databases. Now that the data is in mass-consumption format, organizations can use the Oracle Exalytics In-Memory Machine to deliver a wealth of information to organization analysts. Oracle Exalytics is an engineered system providing speed-of-thought data access for analysts and it is optimized to run Oracle Business Intelligence Enterprise Edition and Oracle Essbase with in-memory aggregation capabilities built into the system.
Step #4: Decide. Oracle's Business Intelligence solutions then provide enhanced capabilities to help organization personnel discover insights from the wide variety of data and create reports using the insights. Oracle R Enterprise, which integrates the open-source statistical environment R with Oracle Database 11g, provides deep statistical analysis for statisticians and data scientists. Oracle's Endeca solutions provide tools for ad-hoc data discovery and exploration by business users. Oracle Real Time Decisions can integrate with big data to automatically take action as certain results are discovered, while the Oracle Business Intelligence Foundation Suite facilitates enterprise-wide reporting and dashboard summaries.

Get More Details
The Oracle Big Data Appliance, along with the Oracle Exadata Database Machine and the Oracle Exalytics In-Memory Machine, provide the core components of a big data solution that can help organizations increase efficiency, better serve constituents, and save money.

CONTACT US
For more information about Oracle's big data solutions go to: http://www.oracle.com/us/technologies/big-data/index.html or call +1.800.ORACLE1 to speak to an Oracle representative.