

**Statement of Darryl McDonald
Of Teradata Corporation
Testimony Before the
House Ways and Means Committee
Subcommittee on Human Resources**

On

April 19, 2012

In

Longworth House Office Building, Room 1100

At

10:00 A.M.

A Hearing Entitled:

**"Use of Technology to Better Target Benefits and Eliminate Waste, Fraud,
and Abuse**

Chairman Davis, Ranking Member Doggett, and distinguished members of the Subcommittee, thank you for the opportunity to testify today on this important subject.

I am Darryl McDonald representing Teradata Corporation. For thirty years, Teradata has continually redefined the leading edge of database technology and advanced analytics, enabling organizations to transform raw data into strategic advantage and attain their visions. As executive vice president of applications, business development and Chief Marketing Officer, I help ensure the innovation behind our products and services meet the evolving needs of our diverse government and commercial customers.

We are living in a time when the volume of data is growing at an enormous rate. Among the nearly 1400 Teradata customers, more than 300 have 100+ terabyte systems while the number of customers with petabyte sized systems doubled between the years 2010 to 2011. To put those metrics into perspective, ten terabytes could hold the printed collection of the Library of Congress; and a petabyte is 1000 terabytes, or the approximate equivalent of 13.3 years of HD-TV video. These customers have seen the advantages of using the data available to them – not by keeping the data in data silos but by integrating it to bring value to the organization - through improved operations, service delivery and customer engagement. The returns on investment from the successful implementation of these systems have been substantial.

The list of companies using Teradata includes technology leaders from across all industries including eBay, Walmart, Caterpillar, Wells Fargo, United Airlines, AT&T and Ford to name a few. These companies are getting smarter by not only recognizing data as their most valuable asset, but also by transforming that data into useful insight for better decision making. They further use this intelligence for analyzing customers' experiences and preferences, to create new products and improve service delivery. We believe that these approaches can be applied by government agencies to streamline operations and improve service delivery while achieving the goals of their missions. Mr. Chairman, we also believe that smarter government is smaller government.

My testimony today focuses on applying advanced analytics to massive data sets – what Teradata terms “big analytics” – to help federal agencies meet complex, large-scale mission demands despite unprecedented budget cuts and constraints.

How *Moneyball* Helps Explain the Value of Big Analytics

It's important to remember that even though big analytics helps address huge, complicated problems, the principles behind it aren't futuristic or incomprehensible.

In fact, one of the best known recent examples that illustrates the value of big analytics doesn't come from government, but from Major League Baseball: the 2002 Oakland A's, as described in Michael Lewis' 2003 book *Moneyball: The Art of Winning an Unfair Game* and the 2011 feature film *Moneyball*.

The book and film relate how the A's' used data analytics to successfully compete better against other teams that could outspend them to attract the best talent. The data analytics used quantitatively analyze player data. Armed with insight provided by the data analytics, the A's assembled a team of players undervalued by traditional baseball criteria.

Though controversial at the time, Oakland's bet on data analytics paid off handsomely and changed the sport forever. The A's finished the 2002 season first in the American League West with an American League record 20-game winning streak that still stands today. Two years later, data analytics helped the Boston Red Sox break the "Curse of the Bambino" and win their first World Series in 86 years. Since then, many teams have added full-time analysts to their staffs.

The parallels between *Moneyball* and our current situation are many, including:

- Just as the A's found value in undervalued players, the federal government has lots of undervalued data that it can, and must, leverage to improve mission success.
- Unable to spend their way out of problems, the A's had to discover what metrics most contributed to mission success and cost-effectively manage legacy and new assets to accomplish that goal. The federal government faces similar challenges today.
- Both the A's in 2002 and the federal government today face entrenched cultural resistance to change from both within their organizations and from their larger communities. But once organizations see the benefits of change – and data analytics in particular – they can move quickly to adopt it and see immediate, widespread and substantial benefits.

Moneyball is an excellent example of how data analytics can transform individual organizations and entire industries. It's especially relevant in this forum because if any pursuit is more devoted to statistics than baseball, it's government. As a nation, we of course have more on the line than winning baseball games. But the principle still stands that leveraging insight from integrated data provides considerable advantages to accomplishing federal agencies' diverse missions.

A Lesser Known Example: Michigan's Data Warehouse Delivers \$1 Million per Business Day

A lesser known example from a Teradata state government customer can be found in the State of Michigan. One of our most successful public sector clients, recognized by the 2012 ComputerWorld Laureate program, is Michigan. Initiated in 1996, the Michigan enterprise data warehouse (EDW) is accessed by 10,000 users in five departments, 20 agencies and 100 bureaus, mostly used within the areas supporting Medicaid/Health and Human Services.

No other state in the United States has achieved Michigan's level of proven results from an EDW. Users include: Department of Community Health (MDCH), Department of Human Services (DHS), Department of Natural Resources, Department of Licensing and Regulator Affairs (DLARA), Secretary of State, Department of Treasury, plus 10 judicial and justice organizations.

As one of the primary users of the EDW, DHS is a perfect example of how the data warehouse demonstrates best practices for a state agency. Traditional data processing systems are absolutely necessary and are the life-blood of DHS. They are used to determine eligibility and issue benefits payments to clients receiving public assistance. However, the agency's data is not accessible in a useful form to analysts and administrators who need it the most. The information is typically entered and stored on many different computer systems, each serving a specific purpose and using its own format. Quite often, lack of consistency makes it extremely difficult to correlate data by drawing on information from several databases. What's more, the level of technical difficulty involved usually makes it impractical for administrators to perform their own interactive queries on the data. The end result is that agencies have become data rich, but have limited ability to transform this data into useable information.

In general, without the EDW, most of the time would be spent looking for data, rather than analyzing it. Because the EDW allows for interactive follow-up questions, it is much more useful than standard, pre-existing reports. The EDW helps uncover trends that previously would have remained hidden.

A few of the most prominent DHS projects include:

- *Food Assistance Program (FAP) Automated Find and Fix (AFF)*—by identifying food assistance cases not receiving correct benefit levels, this project identified and corrected almost \$500 million dollars that were mis-issued to FAP clients from 2005-09.
- *Fraud Detection*—the Office of Inspector General (OIG) uses the data warehouse heavily in detecting fraud within the various program areas that DHS administers. Primarily focused on Child Day Care, Medicaid, Food Assistance, and Cash Assistance, the EDW has assisted OIG in identifying \$70+ million in the last four years.
- *Bridges Information Management Mart*—used to balance caseloads among workers, this online system provides case listings and counts, and provides

performance metrics to help assist in the day-to-day management of the local office. It is estimated to be responsible for \$18 million per year in streamlined labor efficiencies. For example, the amount of time that a worker spends putting together benefit redetermination packages has been reduced by 75%.

- *Child Support*—holds a collection of data from other agencies, which allows a very sophisticated and comprehensive method to “locate” non-custodial parents for the purpose of collecting child support. Ranked as one of the top-performing states in the country, the Office of Child Support is perennially rewarded with high levels of federal incentive money. In addition, these child support collections offset a large amount of money that otherwise, would have been distributed to clients as public assistance dollars.
- *Family Reunification*—DHS and the State Court Administrative Office share data and conduct analytics to protect and care for children in cases involving child abuse, neglect, foster care, adoption and legal guardianship among the 16,000 children enrolled in child welfare services in Michigan. Over nine months after initial implementation, Michigan increased family reunifications by 34% among temporary court wards.

The data warehouse has evolved from its original use to become the primary tool used to bring information from all 12 MDCH separate health-related program areas, encompassing 34 separate data sources, into a single environment. Today, MDCH can monitor the cost and care associated with a single individual across multiple programs. The data warehouse has since become a critical and productive part of the agency’s efforts.

Michigan is a leader in using data integration and yielding a 15:1 cost effectiveness ratio for ONE state agency. Michigan stands as an example for all governments on sharing data across multiple agencies.

Big Data Complicates Decision-Making

Imagine the benefits achieved by Michigan at the scale of the federal government. The federal government faces a broad array of difficult challenges, from ensuring national and homeland security to providing essential services and encouraging U.S. economic competitiveness at home and abroad. A common trend among all these different responsibilities is the massive increase in the amount of data government agencies create and collect as they perform their duties. In that respect, the federal government reflects the global trend of near-exponential data growth over the next decade. IDC estimates that the amount of data worldwide doubles every two years,ⁱ with 1.8 zettabytes – 1.8 trillion gigabytes – created and duplicated in 2011 alone.

Not surprisingly, this blizzard of big data has the potential to overwhelm many federal decision-makers. These massive data sets involve complex and varied data that is both structured, such as spreadsheets; and unstructured, which includes most of the other data organizations today collect, from emails to social media content to video footage. Much of this data is in non-standard formats and resides on non-interoperable legacy systems in multiple locations. This data explosion leads to an obvious question: How can decision-makers at every level of an organization be confident they're drawing the right conclusions from all that data to ensure they best accomplish their agencies' missions?

Getting the answer right to this question is mission critical for federal agencies. The ability of federal decision makers to utilize all data on hand to find underlying insights is essential for effective risk management, cost containment and mission success. This is especially true as the federal government tries to weather record budget deficits and navigate through the worst economy since the Great Depression by mandating agencies find the most cost-effective solutions possible.

What's more, all these trends I've just described are accelerating. In its *2012 Trends to Watch: Government Technology* report, the research and consulting firm Ovum posits that data is replacing oil as the main driver of the global economy, with analytics tools a key means to extracting value from that resource.¹⁴ The omnipresence of the Internet and social media are driving societal demands for government to provide authoritative, accurate data and make fact-based, real-time decisions from it. As one of the world's largest and most influential creators and consumers of data, the federal government will see its long-term success or failure inextricably linked to how it handles big data, not only for research and investigation but for also managing its own business operations.

Big Analytics Solutions Help Overcome Persistent Problems

To overcome the challenges I've just described, federal decision-makers need big data analytics tools that:

- help them extract actionable insights and deliver the right ones to the right people in time to make a difference
- handle multiplying, diverse and overlapping avalanches of structured and unstructured data
- maximize resource efficiency and improve secure collaboration and productivity
- avoid data redundancy and improve data quality and re-use
- increase confidence in the data that is used for decision making
- improve end results for both users and those who rely upon them

Big analytics solutions help assure decision-makers that their important decisions are fact-based and not weakened by data sampling errors because all integrated data is available for analysis. They cut analysis time from weeks or months to near-real time and enable continuous improvement to meet constantly changing technology, regulatory and mission needs.

By helping evaluate past and current data and determine the most cost-effective courses of future action, big analytics solutions are invaluable in finding and fixing the root causes of persistent, confounding problems. Their usefulness is visible across government, from improving cybersecurity and military operations, to tailoring regulations and policy and rooting out fraud, waste and abuse, to understanding the needs of citizens and making a difference at an individual level.

I am honored to share Teradata's role in helping government customers make the best decisions from big data. Teradata is among the world's largest companies focused solely on big data analytics and data warehousing. Our big analytics offerings enable organizations to scrutinize enormous data sets, find actionable insights and gain advantage from those insights most cost-effectively. We empower individuals – government employees, contractors, partners and citizens – to do more, be more economically productive and transform the way government agencies operate.

I would also like to emphasize that Teradata stands ready to help your fellow colleagues and other government officials understand the possibilities big analytics presents. Agencies already have the data needed to deploy big analytics today. Every day the federal government puts off implementing big analytics on the data it already has means money wasted and millions of Americans not receiving essential services more efficiently.

Federal and State Big Analytics Success Stories

I have already mentioned the success achieved in the State of Michigan, but Teradata has nearly 1400 global commercial and government customers worldwide, and we have numerous success stories of helping public sector clients integrate, understand and leverage big data. Teradata's federal customers in the U.S. include the U.S. Department of Health and Human Services Centers for Medicare and Medicaid Services (HHS CMS), the U.S. Air Force Materiel Command, the U.S. Department of Defense Transportation Command (TRANSCOM), Naval Air Systems Command (NAVAIR), and the U.S. Department of Justice, among many others that extend our reach to both the state and national levels of government globally.

Our examples include:

- The U.S. Air Force (USAF) Materiel Command integrates more than 60 source systems to provide clear and accurate decision-making information for the USAF logistics community to quickly provide the ability to see where resources are currently deployed throughout the global supply chain, allowing the best course of action in response to critical time sensitive events. Intangible benefits include access to accurate, reliable, timely and trusted information. The prior process required an Air Force employee to access each and every information system, often logging onto more than 60 systems to run a query, download the results, aggregate the results, then compile the results in an easily understood format. This process could take hours to days, depending on how complex the question was. Today, the same information is integrated into a single system and available in near real-time with a couple of clicks. The speed and accuracy of obtaining this information cannot be precisely measured with metrics or cost savings; rather by customer satisfaction and executing the Air Force mission.
- The US Transportation Command, which provides air, land and sea transportation for the Department of Defense, had previously used a legacy system for supply chain and distribution visibility. This system, however, provided limited visibility to shipment and arrival detail with limited historical data. Today, they have the ability to bring together data from 33 different internal systems plus 600 commercial feeds to improve their decision making. Different agencies and functional areas now access this information from a single source, providing near real-time visibility of cargo and personnel movements, trend analysis and projections on 5 years of data, cost reductions, streamlined operations and improves the logistics information sharing across the Department of Defense.
- Eight U.S. states have recovered \$1.6+ billion – and counting – in recovered tax revenues from non-compliant tax filers across by integrating data from the IRS, other state agencies, and external data. By identifying anomalies in the data, the states target areas to improve taxpayer and tax preparer education, provide data for more informed policy decisions, and more easily identify under- and non- payment of taxes. Efficiencies are further gained by providing access to a single view of the taxpayer audit through a case management system, available to any of the approved staff of the tax and revenue compliance departments, eliminating the need for inefficient paper processes and improving customer satisfaction.
- Another example includes the Centers for Medicare and Medicaid Services (CMS) within the Department of Health and Human Services. As part of the open government initiative, CMS was able to simplify and make its data more accessible to researchers and policymakers through the CMS Dashboards. The initial implementation was completed within 5 weeks – one week short of the deadline. They were able to make this data available in such a short

timeframe because CMS already had claims data available in a single system going back to the year 2006.

- U.S. Department of Agriculture (USDA): Teradata has helped the U.S. Department of Agriculture's Risk Management Agency's Center for Agribusiness Excellence, led by Dr. Bert Little of Tarleton University, see marked success with data analytics to avoid an estimated \$1.5 billion in related costs. The Center use data analytics to ensure the integrity and cost-efficiency of the Federal Crop Insurance program, particularly by combating fraudulent claims. From 2001 to 2011, the Center saved American taxpayers approximately \$838 million in improper payouts.

With many government organizations seeing tremendous benefits from the big analytics deployments they have made so far, a logical next question is how to overcome the challenges I described above to motivate more government agencies to adopt big analytics. Profit, shareholders and competitive advantage motivate commercial users, while regulatory compliance, cost containment and improving mission performance drive government agencies.

In that respect, the timing of this hearing couldn't be better. Legislation requiring standard data content and formats currently has bipartisan support in both chambers of Congress. Data standardization improves the efficiency of big data exchanges, which makes adopting big analytics easier and more cost-effective. Last September, President Obama signed the first such bill into law: the Child and Family Services Improvement and Innovation Act of 2011, which requires data standardization for federal human services programs. In addition, the same provisions have been enacted recently in unemployment insurance and TANF programs (PL 112-96). H.R. 3339, the Standard Data Act, introduced last November builds off that success as well as H.R. 4282 which addresses the needs in the child support program.

Mr. Chairman, thank you for your leadership on this legislation. Teradata strongly supports both H.R. 3339 and H.R. 4282 and respectfully requests Congress to pass this much needed legislation as quickly as possible. Not only would this legislation improve cost-effective delivery of essential services to millions of Americans, but it would also serve also as a stepping stone for other government agencies to understand and acquire the benefits of big analytics. Consider the lessons from *Moneyball*, the best decisions wins and those decisions are based on the ability to integrate data, make it accessible, and perform big analytics. The better agencies know what data they have and how they can use it, the more they can cost-effectively protect our country, improve service delivery and eliminate waste, fraud, and abuse.

Once again, Mr. Chairman, thank you for the opportunity to testify today, and I look forward to your questions. Thank you.

ⁱ<http://www.emc.com/about/news/press/2011/20110628-01.htm>

ⁱⁱ<http://www.itchannelinsight.com/2011/11/cloud-data-analytics-and-agile-development-the-future-of-government-it/>