

PUBLIC SERVICE GETS



BY TAM HARBERT

SMARTER

DATA WAREHOUSING AND
BUSINESS INTELLIGENCE HELP
THE PUBLIC SECTOR SHARE AND
ANALYZE VALUABLE DATA STORES

When it comes to using data warehouses and business intelligence (BI), the public sector has historically lagged the private sector. Part of the reason may be that public organizations face challenges that are greater than—or at least different from—the private sector’s hurdles when implementing these projects.

First, finding the money to pay for improving IT is seldom easy and can be even more difficult for public-service groups amidst today’s budget shortfalls. Second, public institutions often run into substantial barriers when it comes to sharing data. And third, a public organization’s IT strategy can shift rather arbitrarily due to frequent leadership turnover.

For a sector that’s facing new demands for transparency and performance, the benefits of analytic software powered by data warehouses are proving irresistible, especially given that recent legislation requires public institutions to better manage and track the use and effectiveness of federal funding. As public organizations adopt these technologies, they’re finding new solutions to their long-standing challenges.

Data reporting regulations ring the BI bell

Any discussion of data warehousing and BI in the public sector has to start with money—or too often, the lack of it. Against

that backdrop, the American Recovery and Reinvestment Act (ARRA), passed by the U.S. Congress in February 2009, promises to shower US\$787 billion across the country to help promote job growth and economic activity. Twenty-nine federal agencies are tasked with distributing the funds. State and local governments will receive more than US\$541 billion in discretionary and direct spending and tax cuts; of that total, US\$204 billion is specifically tagged as discretionary spending, according to INPUT, a market research firm that tracks government spending. Some US\$100 billion of ARRA funds is going to public schools and colleges, including US\$41 billion in grants to local school districts.

The Mobile County Public School System (MCPSS) in Alabama is one of them. The largest school district in the state, MCPSS has 63,000 students that generate 1.2 million attendance records and 3.5 million grade records a year. This and other data reside in three different databases (human resources, school and student information, and federal programs) in three different physical locations. Gathering the right information to create timely reports was difficult and time-consuming. For example, federal programs want test scores coupled with both student attendance and teacher qualifications, all of which were in different databases.



“We were spending just incredible amounts of time building these reports, pulling the data sources together to build queries, to build analysis tools,” says David Akridge, CIO of MCPSS. “We had to pull all this together in multiple data sets, push it into an Excel spreadsheet or something like that that could be given to them,” he explains. “If they wanted to change something, we’d have to go back and rebuild it.”

Akridge saw the stimulus money program as an opportunity. He convinced the board of education to revive a former attempt to build a data warehouse to help the district pull relevant information together more quickly and efficiently. The board gave its approval and, after extensive review of competitive offerings, the district hired IBM and its local business partner, DecisionEd Group, to develop a data warehouse based on IBM analytics and BI technology. By August 2009, the district had launched the new system and was rolling it out to teachers and administrators.

MCPSS is not unique. Public-sector organizations across the country, from schools to state and local governments, are using stimulus money to embrace data warehousing and BI. “Business intelligence and performance management have always been strong in the public sector, and the stimulus is just making it stronger,” says Robert

Dolan, IBM global government and education industry executive, BI and performance management. “We’re seeing a lot of interest from government organizations that probably didn’t think they needed the technology. There’s a buzz around it right now.”

In fact, the ARRA virtually mandates the use of BI in government by including strict requirements for accountability and transparency in the use of taxpayer dollars (see sidebar, “Stimulating intelligence”). Organizations that receive stimulus funds are required to publish accounting, allocation, and results data for the money received. The law also mandated the creation of a Web site, Recovery.gov, which is intended to provide increasingly detailed information to citizens on how stimulus funds are being used.

Limits to sharing data

But public-sector organizations face unique challenges when implementing data warehouse and BI projects. For example, when a corporation wants to share data, it usually has the option of standardizing on a certain database platform and sharing data freely among its business divisions. That’s a grand oversimplification, and the reality is usually fraught with internal politics and technical challenges, but

Mobile County Public School System

Project: Build a data warehouse to improve information delivery and enable more effective student management

Cost: US\$1.2 million

Challenge: Information was in three different physical databases. Rendering reports was a time-consuming, complex process that provided limited insight into student performance. School administrators had to wait for quarterly reports, which arrived too late in some cases to flag at-risk students before they got into trouble or dropped out of school.

Solution: IBM Cognos Business Intelligence software (ibm.com/cognos), deployed in a relational database environment. Cognos BI produces customizable dashboards that give administrators and teachers up-to-date reports and measures, enabling them to effectively monitor each student based on a number of factors, such as attendance and grades. The system can proactively alert teachers and administrators if a confluence of these factors crosses a threshold, indicating that a student may be at risk.



the path to data sharing in the private sector is—at least theoretically—relatively straightforward.

In the public sector, it can be a lot more complicated. Start with the fact that there may be legal restrictions on sharing data. Even state agencies that send their reports to the same person—the governor—may have different rules and legal restrictions on distributing data, says Geoff DePriest, senior manager in the performance business unit of Crowe Horwath LLP. If one agency is getting funding from the U.S. Department of Labor and another is getting funding from the U.S. Department of Housing and Urban Development, for example, they may have to abide by different rules, he explains.

Even without legal restrictions, effective data sharing is often thwarted because of another obstacle familiar to private-sector businesses: various agencies structure the same data differently. For example, when DePriest worked for the State of Indiana, the Department of Workforce Development architected its data on a person-by-person basis, while Family Social Services structured the data on a case-by-case basis.

One agency that is trying to surmount such data-sharing challenges is the U.S. Census Bureau. In 2007, the Census Bureau granted a contract to IBM to provide data tabulation and dissemination services to support the 2010 Census and other key surveys. The Data Access and Dissemination System (DADS) division of the Census Bureau is responsible for disseminating data from five major surveys, including the decennial (10-year) census that will be conducted in 2010. But each of the five surveys collects data differently, and they use different database systems on the back end to summarize the data. “One shop may be using SAS, another produces database tables, and in some cases we simply get flat files,” says Jeff Sisson,

DADS program manager. Because of this, the IT staff spends a lot of time custom-coding data for the system.

With IBM’s help, DADS is building a data-loading system that will be flexible enough to handle a variety of data formats. By using standard technology and designing the system around metadata, DADS expects to increase efficiency and save back-end costs.

In addition, IBM is coordinating work on American FactFinder (www.factfinder.census.gov), the Census Bureau Web site that publishes data for public consumption, to make it easy for the average citizen to search for and work with data on the Web site.

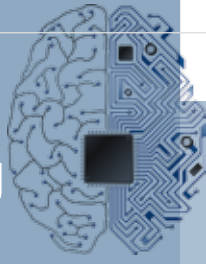
On the current version of American FactFinder, for instance, if citizens want to see all married households in the United States, they must first specify which of the five major surveys to search. “We’re going to make a significant leap forward when we go to the new dissemination system,” says Sisson. “We’re changing the main paradigm in terms of search and navigation.”

Rather than being survey-driven, the revamped site’s search function will be topic-driven. “If I want to find information on married households, I just enter that as a search term and it will bring up everything on married households, regardless of what survey the data is from. It will be a much more powerful tool for Joe Public,” adds Sisson.

The system will also incorporate more sophisticated search and navigation tools, as well as enhanced mapping and charting capabilities. The goal is to accommodate a variety of users with a wide range of skill levels.

The new American FactFinder Web site should be ready for the public by January 2011, says Sisson. And, by law, the data from the 2010 census has to be loaded into the system

Stimulating intelligence



Four parts of the American Recovery and Reinvestment Act (ARRA) are likely to require BI, according to Ramon C. Barquin, president of the Barquin International consultancy and co-founder of The Data Warehousing Institute.

1. Upgrade general government IT.

Significant mandates require funds be used specifically to develop, enhance, or modernize IT systems throughout the federal and state government system.

2. Develop health IT systems.

“Business intelligence will, of course, be one of the pillars of health IT since the massive amounts of data from electronic health records will be the prime object of significant analysis,” according to Barquin.

3. Drive education improvements.

The law sets out specific goals

and reforms, and requires grantees to measure and track progress toward those goals. “A boatload of metrics is mandated from grantees, whether they be states, local governments, contractors, or institutions,” Barquin notes.

4. Require transparency and accountability reporting.

The law requires all recipients to track how funds are used and how many jobs are created, for example. “Without a robust BI toolkit, recipients of ARRA funds will probably not be able to take full advantage of the funding and will not be able to comply with the reporting requirements,” says Barquin.

by March 31, 2011, and distributed to state governments—a challenge that the data-loading system improvements will help address. “That’s all the data that states use to reapportion their congressional districts,” he says. “We can’t be late.”

New election, new boss

Another challenge in the public sector is frequent (and practically guaranteed) leadership turnover. In state governments, for example, priorities can change with every four-year election cycle. “You can’t assume that the next round of leadership is going to want to share and leverage data in the exact same way,” notes DePriest.

U.S. Census Bureau

Project: Improve data tabulation and dissemination for the 2010 Census and other key Census Bureau surveys

Cost: US\$89.5 million

Challenge: Expedite data tabulation assessment for five major surveys of the Census Bureau, including the 2010 Census, the American Community Survey, the economic census, annual economic surveys, and the population estimates program. Increase flexibility in analysis of the data and improve the usability of the information on the American FactFinder Web site.

Solution: IBM Global Business Services is integrating a variety of technologies including a pre-existing data warehouse, IBM WebSphere software, IBM Tivoli Workload Scheduler, Space-Time Research software, an ESRI mapping and charting engine, and Endeca search and navigation solutions.



It happens in school districts, too. MCPSS had been trying to launch a data warehouse project for four years, says Akridge. It put out an RFP back in 2006 and IBM won that bid. Then the district got a new CIO who decided to switch to another company. That project was unsuccessful. Meanwhile, Akridge was appointed CIO and saw an opportunity to try again.

“We’ve been through so much with this,” he says. “We went a long time being very disappointed.” Akridge is thrilled that the project is now going so quickly. “IBM Cognos and DecisionEd have done in four weeks what other companies we’ve worked with couldn’t do in a year. Through invaluable information insights delivered by IBM technology, we’re moving closer to fulfilling our mission of graduating citizens who are prepared with the skills they need for the 21st century.”

As electronic reporting requirements tighten and citizen demands for transparency and access—not to mention the number of data-gathering applications and systems—rise, warehousing, performance management, and BI technologies are likely to become central pieces of the public-sector data management puzzle. An organization may start using these technologies just to track and manage recovery funds or demonstrate compliance with government regulations, but ultimately it can use the systems to make decisions over the long term that improve productivity, reduce costs, and deliver better service. *

Tam Harbert is a Washington, D.C.-based journalist who covers technology, business, and public policy.