Big Data: Strategic Assets

The information states collect can be a force for creating good policy.

BY ALLISON HILTZ

Allison Hiltz is a research analyst in NCSL's State Services Division.

Your phone dings. You pick it up, knowing the familiar sound indicates a new email because you customized the settings yourself. You swipe to open the message and see it's from a constituent, so you quickly respond. But before you close out, you toggle over to view your personal emails. They're all in the same app, after all.

It wasn't so long ago, in the days before smartphones, that you had to log in to and out of individual email accounts and memorize passwords. The accounts were stored separately, requiring independent access. But technology, aided by software development, brought it all together, streamlining your email and, ultimately, your life.

It is a similar integration and improved efficiency that many argue are needed in state governments, particularly when it comes to data.

The Importance of Data

It's no secret that data can help lawmakers improve existing programs and allocate resources. High school graduation rates shape education policies, financial data inform budget decisions and usage statistics guide the delivery of public services. The list goes on.

Used properly, data can improve the effectiveness of state government. Washington's Department of Social and Health Services, for example, used state Medicaid data to target women's tobacco use because it showed that pregnant women enrolled in the program had a higher prevalence of smoking. The result? Healthier mothers and children.

The biggest challenge for many states is not collecting data but figuring out how to use it. Preliminary research by The Pew Charitable Trust's Data as a Strategic Asset initiative found that, while all 50 states incorporate data to some extent to inform decision-making, they all face similar obstacles, including silos, data quality, and security and privacy concerns.

Data as a Strategic Asset

The Pew Charitable Trusts launched the Data as a Strategic Asset project in October 2015 to provide a thorough assessment of the administrative use of data across the 50 states. The two-year initiative is looking at whether and how effectively states are using data for policy, budget and operational decision making. Read case studies at pewtrusts.org, where a full report will be available later this year.

The Challenge of Sharing Silos

Many state government agencies across the country have access only to their own electronic files, creating data "silos"—troves of information that cannot be shared easily across agencies. The left hand sometimes doesn't know what the right hand is doing, creating gaps in services when several agencies are involved in implementing a policy.

Reducing teen pregnancy, for example, might involve the departments of education, health and human services, and child welfare. Similarly, crime prevention might include all those departments, along with mental health and welfare programs. The challenge, then, is to knock the silos down, integrate the data and deliver

more efficient, effective programs.

Why are the silos so hard to break? First, building and sustaining the human and technical capacity to support data sharing and analysis often requires funding investments.

Second, rules often prevent sharing data, such as tax and employment records, among agencies. Adopting collaborative approaches that bring together people from different agencies—and the agencies themselves—to share and integrate data can address this challenge, according to Pew and the Center for Regional Economic Competitiveness.

Data Quality Matters

Bad data can mean big trouble. Maryland learned this lesson in 2007, when a simple typo by the state's Department of Assessments caused one school system to be shortchanged by $24 million, while 17 others were overpaid. Because the schools that received the extra funding were not asked to return the money, the mistake cost the state $31 million.

Last year, Mississippi's Legislative Budget Office identified an accounting error that overestimated general fund revenues by almost $57 million for fiscal year 2017, leading to cuts that resulted in a 1.63 percent funding decrease for most agencies.

Especially challenging are data from surveys, particularly when self-reported. A 2015 study by the National Institutes of Health on self-reported adherence to medi-

cation cited memory biases and discomfort answering sensitive questions as barriers to accurate self-reporting. Even when unintentional, inaccuracies can affect reported health outcomes and, subsequently, funding for health-related policies.

Cybersecurity and Data Privacy

Legislators and cybersecurity professionals differ in their opinions on the effectiveness of current data-security measures, with the latter being less confident than the former that their states are prepared to handle threats. Although 66 percent of legislators say their state is doing enough to address cybersecurity, only 27 percent of chief information security officers do, according to a 2016 report by the National Association of State Chief Information Officers and Deloitte & Touche.

"We are facing a new era of data and information, and with clearer data to make informed decisions come consequences," says Montana Representative Daniel Zolnikov (R). "The more information gathered, the larger the risk of this data being stolen or misused. Privacy is a real concern, and we should be asking ourselves why we are collecting the data, how sensitive is the data and how long should we be holding on to individual information."

Costs are a major impediment to data security. Whereas 80 percent of secu-

Representative
Daniel Zolnikov
Montana

rity officers cited funding as a top challenge in combating cyber threats, according to the Deloitte report, most states spend less than 2 percent of their budgets on cybersecurity.

But without a strong technological infrastructure, the odds of a successful cyberattack increase. A breach can place citizens' personal and financial data at risk and could result in identity theft, financial losses and more.

A Cyclical Data Challenge

To harness data, states need to get buy-in from lawmakers, government departments and state agencies. Once that's secured, technical infrastructure often needs to be upgraded. With new software, hardware, networks or data centers in place, agencies can begin collecting data, but as information accumulates, greater security is required, leading back, of course, to a renewed financial commitment from the legislature.

BUDGETS

Case Studies

Pew identified three states where investing in and harnessing data improved outcomes and savings in postsecondary education enrollment, unemployment insurance and homelessness.

Delaware

Delaware in 2009 became one of 11 states to receive grant funding from the federal Race to the Top program, which aimed, in part, to use data to increase college enrollment. Though Delaware had been collecting longitudinal data on students since 1994, it lacked the analytical capacity to put the information into a usable format. Much of it was stored in several different systems, or data silos, making it difficult to access and analyze.

After receiving the grant and additional personnel through a partnership with the Strategic Data Project at Harvard University's Center for Education Policy Research, the old silo system began to change. The first step was improving the state's technology and data analysis.

Delaware officials decided to use the enhanced data system to address the fact that many "highly qualified" high school seniors (those with SAT scores at 1550 or higher) were not enrolling in college. "Only 30 percent of first-time ninth-grade students completed high school on time, seamlessly enrolled in college the fall following graduation and continued to their second university year," according to Pew. In addition, only 51 percent of recent high school graduates enrolled in college the following fall, compared with a national average of 68 percent. The numbers were even lower for low-income families.

The data were in, but the question remained: What could Delaware officials do about it? Plenty, they decided. They could:

• Improve school counselor training and assistance in financial aid applications.

• Increase applications by first-generation and low-income students through a public education campaign that included then-Governor Jack Markell's proclaiming October and November official College Application Months.

• Develop a text-messaging service to
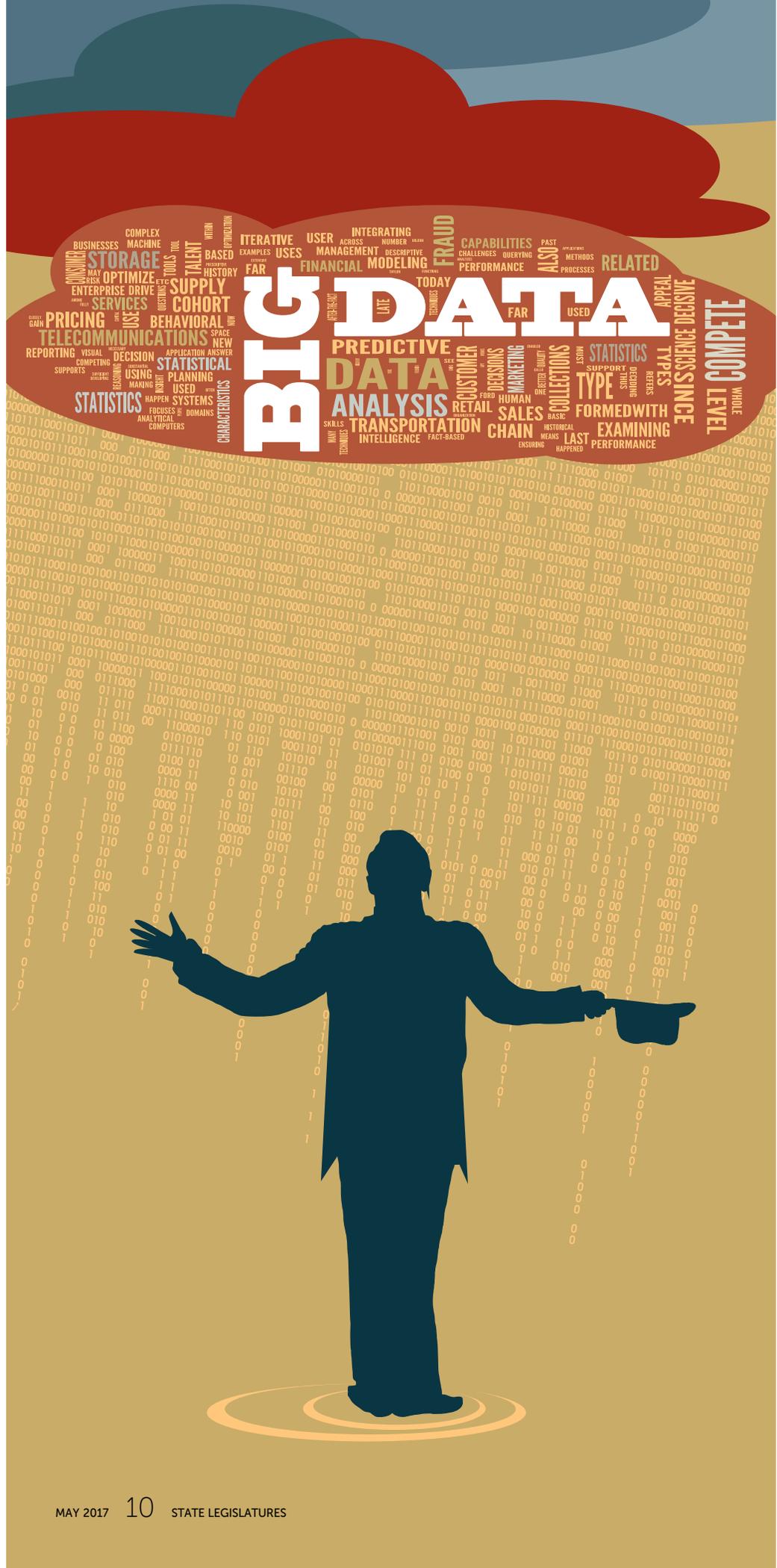
disseminate information and answer questions in real time.

The efforts paid off. In 2015, 100 percent of highly qualified students applied to a postsecondary institution. In 2016, the state expanded its definition of highly qualified to include AP scores, dual enrollment and other factors. Even with the expanded pool, all eligible students applied to a postsecondary institution that year.

New Mexico

New Mexico received a federal grant to help root out overpayments and fraud in unemployment insurance in 2011 and '12. At the time, the state's data, like that in many states, sat in separate silos, with business tax information in one system and unemployment claims and job-separation details in another. Because the systems didn't "talk" to each other, the state could not easily verify that employer and employee claims matched nor compare applicants' Social Security numbers against death records.

Using the grant, New Mexico's Department of Workforce Solutions combined the two systems and set up data-matching capabilities across various agencies, which proved successful. By 2013, the state reduced its unemployment insurance fraud by 60 percent, saving $10 million. Savings leveled off once the initial identifications had been made, so New Mexico took things a step further.

The workforce department partnered with Deloitte Consulting LLP in 2015 to use other data, including work history and prior claims, to curb potential fraud associated with the online unemployment insurance application. Because inaccurate answers to questions on the form—intentional or not—were costing the state money, the focus was on ensuring accuracy.

Deloitte designed a system that uses previously collected data to identify applicants' potentially inaccurate responses. Throughout the process, the system displays a pop-up message, known as a "behavioral nudge," encouraging the applicant to be truthful. Testing showed that personalized messages, such as how fellow county residents responded, were far more effective than generic messages containing threats about laws and penalties.

Equipped with better information, New Mexico reduced its benefit-year earnings overpayments (when someone returns to work but still claims unemployment benefits) by 2 percentage points, more than double the national average. The system also yielded more precise information about claimants' reasons for not working, thus reducing unqualified payments.

New Mexico is now looking for other areas where harnessing data can realize similar benefits, such as health and human services.

Virginia

Concerned about homelessness, then-Governor Bob McDonnell (R) brought together multiple departments and agencies in 2010, and in 2014 newly elected Governor Terry McAuliffe (D) decided to focus on veterans. At the time, veterans made up 10 percent of the adult population but represented 16 percent of the homeless population.

The state analyzed its data, first, to learn about each homeless veteran. Officials then collaborated across departments and agencies to customize services based on their individual needs. The integrated effort allowed them to understand why they became homeless in the first place and to target solutions accordingly.

By 2015, Virginia's homeless veteran rate was the lowest in the nation.

"Data integration is key to ensuring that we can identify veterans experiencing homelessness and quickly move them into permanent housing," says Senator Jennifer McClellan (D). "Sharing data across providers and systems has allowed us to better assess the needs in the Commonwealth, communicate more effectively with partners, and target and prioritize resources to ensure veteran homelessness is rare, brief and nonrecurring."

Senator Jennifer McClellan Virginia

Bolstered by their success, officials are working to prevent others from becoming homeless by integrating and analyzing data from jails, hospitals and law enforcement agencies.

Where to From Here?

Delaware, New Mexico and Virginia, of course, are not the only states using data in their decision-making processes.

Twenty-three states integrate data into budgeting decisions through the Results First Initiative, another program created by Pew and the John D. and Catherine T. MacArthur Foundation.

As states harness the power of their data, program administration is likely to change—and it won't always be easy, something Jonathan Ball, director of Utah's Legislative Fiscal Analysts Office, has witnessed firsthand. Earlier this year, Utah lawmakers opted to re-evaluate some tax reform proposals. "While we did a ton of work," Ball says, "it didn't do what it needed to do, so we needed to take a step back."

Jonathan Ball Legislative Fiscal Analysts Office Utah

But overcoming challenges is something states are equipped to do. And despite the time, the money and, for some, the culture shift required, many are charging ahead.