

COSTS OF STATE LONGITUDINAL DATA SYSTEMS

BY JENNA LEVENTOFF



State Longitudinal Data Systems (SLDS) match data about individuals from different sources over time, and serve as an important resource for state policymakers, researchers, and the public. When states and other entities consider building longitudinal data systems, it is helpful to understand the costs involved. This fact sheet provides examples of select costs associated with implementing and maintaining an SLDS.

The cost of building these systems can vary dramatically. Some states have built new systems for about \$2.5 million while others have spent upwards of \$7 million. Factors influencing startup costs include not just the type of hardware and software purchased, but also the number of participating agencies and amount of data, and how states execute their data governance and policy procedures. Whether an SLDS is “centralized” or “federated,” will also have a bearing on costs. Centralized data systems collect and maintain data in a single location, whereas federated systems temporarily link data for a specific purpose, but house data in individual state agencies. Each system type has its pros and cons. Centralized systems can make it easier to gather and utilize data while federated systems can allow for tailored security protections. Centralized systems tend to be more expensive to build than federated systems but may require fewer resources from participating agencies.¹

Moreover, the costs of an SLDS do not end after the system is built. Like startup costs, the costs of keeping the system in working order, and for using the system to create actionable information for stakeholders, can vary dramatically based upon the analytical capacity needed, how much data is shared and how sophisticated the technology is.

Implementation

PRIMARY DETERMINANTS OF STARTUP COSTS	OTHER COST CONSIDERATIONS
Type of SLDS (centralized or federated)	Vendor or in-house development
Capabilities	Available funding
Previously existing linkages and infrastructure	When the system was built (many technology costs decrease over time)
Number of participating agencies and amount of data	Level of organization and planning
Ease of negotiating data sharing agreements	Complexity of data governance and policy procedures
	Data quality

Annual maintenance

PRIMARY DETERMINANTS OF MAINTENANCE COSTS	OTHER COST CONSIDERATIONS
Amount of data analysis and data products needed	Level of technological sophistication of hardware
Amount of data requests	Software used
Hosting expenses	Amount of data and new data acquisitions
Automated or human research capacity	

Costs of centralized data systems

MARYLAND: Maryland officials estimate that implementation of the [Maryland Longitudinal Data System \(MLDS\)](#) cost \$2,747,00. The Maryland system is a centralized model built in 2014. It primarily contains data from three state agencies: the Maryland Higher Education Commission, the Maryland State Department of Education, and the Maryland Department of Labor, Licensing, and Regulation. Maryland spent \$1.6 million of the implementation costs on hardware and software, including purchasing servers, data center switches, and various software licenses (including Oracle and the MFT license required to securely transfer files from participating agencies). The state paid contractors about

¹ National Center for Education Statistics, Institute of Education Sciences, “Centralized vs. Federated: State Approaches to P-20W Data Systems,” https://nces.ed.gov/programs/slds/pdf/federated_centralized_print.pdf.



**WORKFORCE
DATA QUALITY
CAMPAIGN**

www.workforcedqc.org



NATIONAL SKILLS COALITION
Every worker. Every industry. A strong economy.

\$800,000 to design the system, its architecture and the MLDS website. Finally, Maryland spent just over \$300,000 to hire key staff, such as a chief information officer, a data analyst, and a developer.

The MLDS is managed by the MLDS Center, which is an independent unit of the state government. The agency's annual budget is \$2,077,000. The majority of the funds (\$1.3 million) are for salaries of the thirteen fulltime employees of the Center. Of the employees, eight are information technology specialists who maintain the system in-house, maintain a public facing website with data dashboards and reports, support research and reporting, and fulfill data requests. The staff are assisted by one contractor, who provides senior database engineering services (\$236,000). In addition to salaries, funds are used for hardware and software maintenance, general office management, and to fund a research partnership with the University of Maryland School of Social Work and College of Education (\$365,000).

UTAH: Utah officials estimate that the state spent \$7,144,934 to build the Utah Data Alliance (UDA), Utah's previous centralized longitudinal data system.² This system linked data from four Utah agencies: the Utah State Board of Education, the Utah System of Higher Education, the Department of Workforce Services, and the Utah System of Technical Colleges. Approximately \$500,000 of the implementation costs went towards purchasing hardware and software for the system. The state paid a contractor \$2.4 million to develop UDA and spent the final \$4.2 million on staff salaries and equipment.

Utah reported spending about \$1.8 million a year to maintain UDA. Of that, the state spent approximately \$645,000 to hire staff to maintain the technical system, and pay for hardware and software contracts. The state also spent \$345,000 to conduct research and analysis, and an additional \$200,000 for UDA staff to manage the system. The remaining amount — approximately \$600,000 — was provided to participating agencies to help them pay for the staff and equipment they need to prepare and validate their own data, as well as conduct periodic research requested by the legislature.

Costs of federated data systems

VIRGINIA: Virginia reports having spent approximately \$7.5 million to plan and build the [Virginia Longitudinal Data System \(VLDS\)](#) between 2010 and 2013. VLDS is a federated system which provides access to linked data from nine state agencies, to include Virginia's Department of Education, the State Council of Higher Education for Virginia, Virginia's Employment Commission, Virginia's Department of Social Services, Virginia's Community College System, the Virginia Department for Aging and Rehabilitative Services, Virginia's Department of Health Professions, Office of Children's Services, and the Department of Juvenile Justice. In the initial development phases, the state spent roughly \$6,750,000 on software development and integration services, \$450,000 for staff services, \$300,000 for hardware, and \$75,000 for software.

Virginia spends roughly \$475,000 a year on operations and maintenance costs, including \$325,000 on vendor support, \$100,000 on hardware and hosting, and \$50,000 on software licensing. Operations and maintenance services are provided by a private contractor who provides shared support to Virginia and Nevada.

NEVADA: Nevada officials reported spending approximately \$2.5 million over the course of fourteen months to develop the [Nevada P-20 to Workforce Research Data System \(NPWR\)](#). NPWR is a federated system that links data from the Nevada Department of Education, the Nevada System of Higher Education, and the Department of Employment, Training and Rehabilitation. Implementation costs included roughly \$60,000 in hardware, \$75,000 for software licenses, \$400,000 in agency resources to develop the data sets, and \$2,000,000 to integrate agency software. NPWR was developed using the software from Virginia, thus enabling the state to save about \$5 million as compared to Virginia's initial implementation cost.

Nevada spends roughly \$450,000 a year on the annual operations and maintenance costs that include hardware, software, hosting, and vendor expenses. Maintenance services are provided by the private contractor who supports Virginia and Nevada. In addition, Nevada budgets nearly \$250,000 annually to create or update automated data dashboards. More traditional reports cost approximately \$10,000-\$15,000 to design, develop, test, and deploy, depending upon their complexity.

ILLINOIS: The [Illinois Longitudinal Data System \(ILDS\)](#) is a federated system containing information from the Illinois Board of Higher Education, the Illinois Community College Board, the Illinois Department of Commerce and Economic Opportunity, the Illinois Department of Employment Security, the Illinois Department of Human Services, the Illinois State Board of Education, and the Illinois Student Assistance Commission. Although WDQC could not find ILDS implementation costs, the state reports budgeting \$310,000 a year to maintain the system. Of that, \$250,000 is for staffing support, equipment, software, system maintenance and upgrades, and for training staff to use the system. The other \$60,000 is used to support staff for the state's Governing Board and the two committees that support the Governing Board's work.³

2 Pursuant to new legislation, Utah replaced the Utah Data Alliance (UDA) with the Utah Data Research Center (UDRC). Learn more at: <https://www.nationalskillscoalition.org/news/blog/utah-bill-bolsters-workforce-data>.

3 Illinois Longitudinal Data System, "2016 Annual Report & Plan," May 2016, <https://www.illinoisworknet.com/ILDS/Documents/2016%20ILDS%20Annual%20Report%20Plan.FINAL%205.27.16.pdf>.