

SLDS Issue Brief: Sources and Linking Strategies for Employment Data

SLDS ISSUE BRIEF

January 2023

U.S. DEPARTMENT OF EDUCATION

A Publication of the National Center for Education Statistics at IES

This product of the Institute of Education Sciences (IES) Statewide Longitudinal Data Systems (SLDS) Grant Program was developed with the help of knowledgeable staff from state education agencies and partner organizations. The information presented does not necessarily represent the opinions of the IES SLDS Grant Program.

For more information on the IES SLDS Grant Program or for support with system development, please visit <http://nces.ed.gov/programs/SLDS>.

CONTRIBUTORS

2016 Version

Jeff Sellers, *SLDS Grant Program State Support Team*
Charles McGrew, *SLDS Grant Program*
Members of the 2016 SLDS Data Linking Workgroup

2022 Version

Kathy Gosa and Jeff Sellers, *SLDS Grant Program State Support Team*
Molly Abend, *Maryland Longitudinal Data System Center*
Linda Borkosky, *Kentucky Center for Statistics*
Dana Brandt, *DataSpark, University of Rhode Island*
Emily Cataldi and Jason Crowley, *Iowa Department of Education*
Tod Massa, *Virginia State Council of Higher Education*
Michael McGroarty, *Michigan Center for Educational Performance and Information*
Ronald Parrish, and Jamie Williams, *Virginia Department of Education*
Sam Ribnick, *Massachusetts Department of Elementary and Secondary Education*
Ben Tate, *Oregon Higher Education Coordinating Commission*

An increasing number of states are pursuing opportunities to link education information to data related to the workforce. Longitudinal records based on these data can provide a clearer picture of how well schools, higher education institutions, and workforce training programs are preparing individuals for successful careers, and they can contribute to the development of employment and wage metrics. They also can provide insight into general workforce characteristics, the alignment between workforce needs and the education pipeline, opportunities for economic and workforce development, and the migration of graduates outside the state.

Nationwide, the increased focus on career readiness and return on investment for education and training underscores the need for comparable outcomes

information across states. In the 2021 Statewide Longitudinal Data System (SLDS) Survey, 29 states indicated that they are “operational” in linking K12 and workforce data. A total of 43 states have received grants through the U.S. Department of Labor’s Workforce Data Quality Initiative (WDQI) to integrate workforce and employment data into longitudinal data systems. In addition to supporting state interests in education and employment outcomes, these data efforts feed outcomes-based reporting requirements under the federal Workforce Innovation and Opportunity Act (WIOA), passed in 2014.

As the demand to connect education and employment data grows, states are navigating the challenges of locating essential data across a number of sources as well as establishing agreements and technical

processes for connecting and managing the data. This brief describes common sources of workforce data and processes that states use to link data across education and workforce programs.

Connecting Education and Employment Data

There are different models for connecting education and employment data. In some cases, the state may have a centralized SLDS where copies of data from multiple agencies such as K12 education agencies, postsecondary institutions, and workforce programs are integrated together. Other states may utilize a federated approach where data from the various agencies are connected as needed to produce a report or analysis without maintaining permanent copies of the connected data. The model a state chooses will depend on its unique priorities and circumstances. Regardless of the SLDS model used, the goal for these efforts is to create information from data that will help stakeholders better understand the connections and flow between education and the workforce.

Sources of Employment Data

In the past, secondary and postsecondary education institutions often relied on alumni surveys, follow-up calls, or similar piecemeal methods of collecting employment information from former students. Such self-reported information often is not reliable or comprehensive, and collecting it usually requires significant staff time and resources. Additionally, different collection methods and definitions make it difficult to compare information across schools and institutions, preventing users from seeing their graduates as part of a broad talent pipeline. By linking to or incorporating employment data with data from the education sector over time, P-20W+ (early childhood through workforce) SLDSs can offer more complete, comprehensive longitudinal data than have been available in the past to a wider range of stakeholders.

Most states rely primarily on unemployment insurance wage records for employment data, but some states use additional data sources for information on workers not included in these records. **TABLE 1** on pages 3 and 4 lists workforce data sources that states might include in their SLDSs, as well as limitations and important considerations for using data from each source.

Starting point: unemployment insurance

The Social Security Act at §1137(a)(3) requires employers to report quarterly wage data including worker names, Social Security numbers, and quarterly

wages. In most states, these data are reported to the agency that administers unemployment compensation programs. These state unemployment insurance quarterly wage records constitute the largest share of the workforce data included in many SLDSs. The records include employer-reported wage information for the majority of workers in a state, as well as employer names, addresses, and industries as represented by their North American Industry Classification System (NAICS) codes. They generally do not include information about individuals' occupations or specific worksites in cases where employers have multiple offices. A few significant groups of workers are not covered by a state's unemployment insurance records, including

- federal and military employees;
- many self-employed workers; and
- individuals who live in one state but work in another.

Processes for and limitations on sharing, storing, and using unemployment insurance records are dictated by federal regulations, state mandates, and the terms of data sharing agreements approved by the SLDS partner agencies.

Closing the gap: other workforce data sources

Unemployment insurance records offer a limited amount of information about the majority of a state's workforce, but many states turn to additional data sources for more detailed information about worker occupations and categories of workers that are not reported to unemployment insurance programs. States such as Kentucky are working to leverage their state's department of revenue to determine further information. Income tax records from the state's department of revenue can provide wage information for self-employed individuals, contractors, part-time workers, and employees of businesses that do not report unemployment insurance information. Some states such as Virginia draw job titles, education history, and additional data about employers from statewide employment assistance websites such as those established under the federal Wagner-Peyser labor exchange initiative. Maryland recently established a data sharing agreement with the Maryland Comptroller to receive aggregate, de-identified income information from tax filings.

The State Wage Interchange System (SWIS) is a voluntary program that states may participate in, and which can provide employment data on individuals who have moved out of state or who live in one state but work in another. In addition, some states have established interstate data sharing agreements that may provide information on out-of-state workers. However, many of these exchanges have limits on how

TABLE 1. Types and sources of employment data commonly included in SLDSs

Record Type	Typical Data	Linking Fields	Strengths	Limitations
<p>State unemployment insurance wage records</p> <p>Source: State unemployment insurance programs (often located in state departments of labor, commerce, or revenue)</p>	<ul style="list-style-type: none"> Quarterly wages Industry of employer (NAICS code) Employer name and location (typically the central office) 	<ul style="list-style-type: none"> Social Security number Partial first and last names 	<ul style="list-style-type: none"> Records cover the vast majority of workers. Quarterly data allow for within-year trend analysis. Records are relatively easy to obtain and are the most commonly used source of employment data. 	<p>Records do not cover some self-employed workers, military or federal employees, or employees who live in one state but work in another. Records also do not identify individuals' occupations.</p>
<p>State unemployment insurance claims payments</p> <p>Source: State unemployment insurance programs (often located in state departments of labor, commerce, or revenue)</p>	<ul style="list-style-type: none"> Biweekly payment amount Date of payment Amount of income recipient has earned outside of unemployment insurance Recipient address 	<ul style="list-style-type: none"> Social Security number First, middle, and last names Date of birth Race/ethnicity Gender 	<ul style="list-style-type: none"> Records provide information about individuals currently seeking work. Records cover actual payments made. Social Security numbers are very reliable. Dates of payment allow for calculation of timeframe of unemployment. 	<p>Individuals who pay child support may have two records for each payment period—one showing payment made directly to the state division of child support and one showing payment to the individual. These two amounts need to be matched by Social Security number and added together for a total payment.</p>
<p>State income tax</p> <p>Source: State department of revenue</p>	<ul style="list-style-type: none"> Number of dependents Address Marital status 	<ul style="list-style-type: none"> Social Security number First and last names Date of birth 	<ul style="list-style-type: none"> Records typically cover all residents with income and some without income. Records include address of residence. Records include individuals who live in the state but work in another state. 	<p>There may be strong restrictions on allowable uses of the data.</p> <p>Records also do not identify individuals' occupations.</p>

Record Type	Typical Data	Linking Fields	Strengths	Limitations
<p>Out-of-state and additional in-state employment</p> <p>Source: State Wage Interchange System (SWIS)</p>	<ul style="list-style-type: none"> Quarterly wages Industry of employer (NAICS code) Employer name and location (typically the central office) 	<ul style="list-style-type: none"> Social Security number 	Records provide data for out-of-state workers not included in other state agency sources.	<p>SWIS is designed to support performance reporting for federal and state workforce, education, and training programs, and data must be reported in the aggregate. Use of SWIS data for other purposes may be limited.</p> <p>State participation is voluntary, and data are not available from all states.</p>

their data can be used and use of SWIS data is limited to specific reporting purposes; therefore many states find it challenging to report on individuals who have moved out-of-state.

Measuring additional outcomes

Pursuing data sources outside of traditional education and workforce agencies can provide states with additional context for education and employment outcomes. Data from public assistance services such as the Supplemental Nutrition Assistance Program (SNAP) and Temporary Assistance for Needy Families (TANF), as well as incarceration and probation data from departments of corrections, can help further illustrate connections among education attainment, workforce and human services, and labor market outcomes for individuals and families.

Linking Employment and Wage Data With Education and Other SLDS Data

State processes for linking wage records from workforce agencies with data from other SLDS partner organizations depend on the data available to identify individuals across those sources. Typically, states must develop solutions for two broad matching challenges: linking postsecondary education records to workforce records and linking K12 education records to workforce records.

Postsecondary-to-workforce matching

Linking postsecondary and workforce records for individuals is made easier by the fact that both sectors generally collect Social Security numbers (SSNs). SSNs are the primary identifier used by most state workforce agencies. Colleges and universities may use institution-specific identifiers internally but also collect SSNs for students. SSNs also are key identifiers for postsecondary financial aid services.

K12-to-workforce matching

Rather than SSNs, most state education agencies rely on combinations of data such as name, date of birth, gender, and usually a statewide unique K12 identifier to distinguish individual students. These identifying data have little overlap with employment and wage records, which generally include only SSNs and partial names for individuals. In Massachusetts, the Department of Elementary and Secondary Education (DESE) sends student data to the Department of Unemployment Assistance, which matches the data primarily on first name, last name, and date of birth and scrubs any identifying information before returning matched data to DESE. Like other state agencies, DESE also spends time fostering relationships to ensure that the matching process goes smoothly. Virginia also does not use SSNs to complete K12 to workforce matching, instead relying on

probabilistic matching completed by the Virginia State Council of Higher Education (<https://www.schev.edu/>).

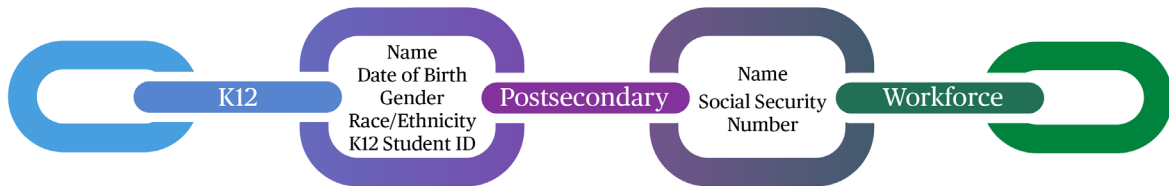
States can use postsecondary education records as a bridge between K12 and workforce data by matching K12 and postsecondary records using demographic data and education identifiers, and then matching postsecondary and workforce records using SSNs and extending that link back to K12 (FIGURE 1). However, this approach covers only individuals who have enrolled in postsecondary programs within the state, which may be a minority of workers in many areas. These limited matches also likely fail to provide information about some of the most vulnerable populations of workers and families who may need the most assistance.

Since few states collect SSNs at the K12 level, many are identifying intermediary sources to help link K12 and

workforce records. Records of driver's licenses and IDs issued by state divisions of motor vehicles (DMVs) are becoming popular sources of matching data. DMV records generally contain individuals' names, dates of birth, and genders—which can be matched to K12 education records—as well as SSNs that can be linked to employment and wage data. States including Iowa, Kentucky, Maryland, Michigan, and Rhode Island have agreements between their respective SLDS programs and DMV or Department of Transportation (DOT) to connect the data using SSN information. The Kentucky Center for Statistics (KYSTATS) also uses driver's license records to improve geolocation data, as these records generally contain up-to-date address information. These data help KYSTATS to determine county of residence, which allows it to report on attributes such as the regions where recent postsecondary graduates are living.

FIGURE 1. K12 education records that do not contain Social Security numbers must be linked to workforce records using intermediary data sources. This figure shows the data elements in postsecondary education and division of motor vehicles (DMV) records that are commonly used to link K12 student records to workforce records.

Linking K12 to Workforce Through Postsecondary Records



Linking K12 to Workforce Through DMV Records



Due to the sensitivity of personally identifiable data, especially SSNs, any process for linking records across multiple agencies must conform to state and federal privacy regulations and include measures to exchange, match, and store data securely. Sometimes this process can be onerous. DataSpark, which manages and staffs the Rhode Island Statewide Longitudinal Data System, entered a data sharing agreement with the Rhode Island Department of Revenue's Division of Motor Vehicles (RIDMV) to measure non-degree credential attainment as required by WIOA and Perkins Act performance reporting. RIDMV interpreted the agreement conservatively. It needed to approve every project conducted using the Rhode Island Statewide Longitudinal Data System and reviewed every product, not just those projects and products using RIDMV data. DataSpark felt that this requirement was too burdensome to continue and destroyed the data when the original purpose was fulfilled and the agreement expired.

Best Practices and Future Opportunities

Collaborating across agencies

Incorporating any new data source into an SLDS requires significant efforts to engage, communicate, and establish partnerships with contributing agencies. The guiding vision, authority, and procedures for sharing and linking data across SLDS partners need to be formally established in a memorandum of understanding (MOU) or in data sharing agreements signed by agency leaders.

Iowa Workforce Development maintains a relationship with the Iowa DOT through an MOU to access DOT records, which contain SSNs and are used as an intermediary to link K12 and workforce records. Data sharing between the Maryland Longitudinal Data System Center and the Maryland Motor Vehicle Administration (MVA) is facilitated through data sharing agreements and state law permitting MVA to disclose data to another state agency to carry out its official duties. To guarantee appropriate security measures and create buy-in from data sharing partners, KYSTATS requires MOUs as well as additional agreements with approval signatures from each data contributor to release data for evaluation or research purposes to external researchers

Even with legislative authorization and data sharing agreements in place, high-quality and timely employment data cannot be obtained without buy-in and cooperation at all levels of partner organizations, from executive leaders to data stewards. Strong cross-agency data governance groups and processes should engage key data managers and decisionmakers

to ensure appropriate handling and use of the data. Additionally, education agencies should work with their workforce partner agencies to identify and communicate benefits to the citizens, agencies, and the state from linking their data to the SLDS, how this collaboration will help education and workforce partners across the state, and how their data will be safeguarded and used appropriately.

Using linked data

When linked to education records, individual employment data provide states with invaluable insight into the impact and value of education and training programs. Federal initiatives such as WIOA and state interests in economic development and return on investment emphasize employment rates and wages as key measures of the value of such programs. However, even though job prospects and potential earnings are important and measurable, they are not the only factors in individuals' education and training decisions, nor do they represent the only value of education.

By making employment outcomes information available, states can work toward identifying and providing additional data to calculate more comprehensive indicators of return on investment for education and training programs. They also can provide useful consumer information to students, workers, and families as they decide how to continue their education to meet their goals.

For example, as a result of linking education data to workforce records, KYSTATS can examine outcomes for postsecondary graduates beyond degree attainment. KYSTATS' December 2021 Postsecondary Feedback Report drew on SLDS data to determine that 57.5 percent of Kentucky postsecondary graduates were employed within Kentucky 3 years after graduating. This report allows users to see outcome data by specific major group, student origins, and transfer students. In Massachusetts, DESE uses linked K12 and workforce data to show earnings and employment outcomes through a public dashboard. Michigan's CEPI uses its linked data to produce a variety of workforce reports, including Entry Level Wages Over Time, Median Wages by Educational Attainment, Median Wages by Field of Study, and more. The data also support the functionality of the Michigan Pathfinder application (<https://pathfinder.mitalent.org/>), which allows prospective postsecondary students to explore careers, schools and programs within Michigan institutions and see the outcome data associated with different selections. DataSpark offers reports exploring Rhode Island's healthcare and social assistance workforce, non-

State Examples: Relationship Building

The Oregon Longitudinal Data Collaborative (OLDC) partners with both the Oregon DMV and the Oregon Department of Human Services (ODHS) to acquire SSN data. Oregon state law allows the OLDC to request SSNs from its partners as a public records request. The OLDC was able to expedite and smooth this process by building trust and transparency through meetings with DMV and ODHS leaders to share the purpose of the longitudinal program and how data are used, stored, and made secure.

The Michigan Center for Educational Performance and Information (CEPI) also credits relationships and trust-building to successful data sharing agreements. Although it was challenging at first to build relationships with each contributing agency and ensure that all parties understood the work that needed to be completed, CEPI fostered trust by adhering to data sharing agreements and not using SSN data outside of agreed-upon projects without first amending agreements and getting approval from the data owners.

The COVID-19 pandemic brought many challenges to data collection and research. In Massachusetts, these challenges increased the workloads of staff and contributed to a great deal of turnover in the Massachusetts Department of Unemployment Assistance (DUA), which provides workforce data and matches them to K12 data for the Massachusetts Department of Elementary and Secondary Education (DESE). These challenges made it difficult for DUA to prioritize data matching. To overcome this issue, DESE staff have focused on building better personal relationships with DUA and having a prep call before each match file is sent.

degree credentials, wages of high school graduates, and employment outcomes of a nursing-focused charter school.

The Maryland Longitudinal Data System Center produces an annual report on the workforce outcomes of high school graduates 5 years after graduation. Each report contains a special supplement focused on outcomes such as college enrollment patterns, college affordability, wage variations by demographic or economic group, and other topics. (<https://mldscenter.maryland.gov/CenterReports.html>) Additionally, Maryland produces numerous dashboards, one of which is College and Workforce Outcomes for Maryland Public High School Graduates (<https://mldscenter.maryland.gov/dashboards.html>).

Although these data are valuable, agencies using linked data must use discretion, particularly in terms of privacy and safety. These data must be reported in the aggregate to the degree that no individual performance or outcome can be identified.

Incorporating additional data sources

Increasing interest in workforce outcomes can both highlight the limitations of existing SLDS workforce data and generate support for incorporating data from additional sources. The gaps in employment data from unemployment insurance records and the need for more comprehensive data have led some states to open discussions about data sharing with organizations not previously involved in SLDS work, such as Social Security Administration and revenue offices, and the Department of Defense for military records.

Expanding data collections and improving data quality

As states use employment data for federal and state reporting, policy analysis, and research, they also identify opportunities to improve existing data collections to produce more accurate and reliable information. For example, collecting complete names and dates of birth in unemployment insurance records would make matching them to K12 education data easier and enhance the quality of those matches.* Uniformly storing K12 student identifiers in postsecondary education records—or establishing a statewide unique identifier to take the place of SSNs in both education and workforce records—also would streamline data matching across agencies. Expanding wage record data collections to include occupations, hourly earnings, and other employment details could provide greater insight into education and employment outcomes. Using employment data for reports and research also will help draw attention to data quality issues in specific datasets and pinpoint areas for improvement.

In most states, plans to alter or expand existing data collections will face numerous hurdles, including compliance with state and federal regulations, legal review, approval by state leaders, and implementation of administrative and technical processes to collect the data. Even so, demonstrating the value of comprehensive workforce data to state and federal

* Some states indicate that they are considering JEDx (<https://www.uschamberfoundation.org/JEDx>) as an infrastructure to collect unemployment insurance records and other jobs data.

policy objectives can create a solid starting point for these discussions.

Additional Resources

SLDS Issue Brief: Addressing Employment Outcome and Workforce Questions
<https://slds.ed.gov/#communities/pdc/documents/17952>

SLDS Issue Brief: Best Practices for Calculating Employment and Earnings Metrics
<https://slds.ed.gov/#communities/pdc/documents/13014>

SLDS Issue Brief: Evaluating the Effectiveness of Matching Methodologies
<https://slds.ed.gov/#communities/pdc/documents/20388>

SLDS Issue Brief: Interagency Data Linking and Common Identifiers
<https://slds.ed.gov/#communities/pdc/documents/19047>

SLDS Webinar: Employment Outcome Indicators
<https://slds.ed.gov/#communities/pdc/documents/9944>

SLDS Webinar: The Match Rate Dilemma
<https://slds.ed.gov/#communities/pdc/documents/8982>

Workforce Data Quality Campaign (WDQC) State Solutions
<http://www.workforcedqc.org/state-solutions>

U.S. Department of Labor: State Wage Interchange System (SWIS)
<https://www.dol.gov/agencies/eta/performance/swis>

State Resources

Kentucky

Kentucky Center for Statistics
<https://kystats.ky.gov/>

Kentucky Postsecondary Feedback Report
https://kystats.ky.gov/Reports/Tableau/2021_PSFRR

Rhode Island

DataSpark, University of Rhode Island
<https://datasparkri.org/>

Identifying Credentials of Value in Rhode Island
<https://datasparkri.org/credentials-of-value>

Rhode Island's Healthcare and Social Assistance Workforce Report
<https://datasparkri.org/healthcare-workforce>

Iowa

Iowa State Longitudinal Data System
<https://educateiowa.gov/data-reporting/state-longitudinal-data-system>

Maryland

Maryland Longitudinal Data System Center
<https://mldscenter.maryland.gov/>

Maryland Longitudinal Data System Center Dashboards
<https://mldscenter.maryland.gov/dashboards.html>

Maryland Longitudinal Data System Center Reports
<https://mldscenter.maryland.gov/CenterReports.html>

Massachusetts

Massachusetts Department of Elementary and Secondary Education
<https://www.doe.mass.edu/>

Massachusetts Employment and Earnings of High School Graduates
<https://app.powerbigov.us/w?r=eyJrIjoiODAzNDU1NWMTOTJiMS00YTZiLTg1MmEtN2JhOTM4ZjNmNWExIiwidCI6IjNlODYxZDE2LTQ4YjctNGEwZS05ODA2LTJhMDRkODFiN2IyYSJ9>

Michigan

Michigan Center for Educational Performance and Information
<https://www.michigan.gov/cepi>

MISchool Workforce Landing Page
<https://www.mischooldata.org/workforce-landing-page/>

Oregon

Oregon Longitudinal Data Collaborative
<https://www.oregon.gov/highered/research/Pages/OLDC.aspx>

Virginia

Virginia Longitudinal Data System
<https://vlds.virginia.gov/>