The Roadmap to a North Carolina Longitudinal Data System (NCLDS)

Christen Holly, Research Assistant Professor (EPIC)
Zach Ambrose, Consultant, Ambrose Strategy
Julie Marks, Senior Research Associate and Director of Program Evaluation (EPIC)
Daniel Gitterman, *Duncan '09 and Rebecca Kyle MacRae* Chair and Professor of Public Policy
Rachel Rana, Research Specialist (EPIC)
Dillan Bono-Lunn, Research Assistant

June 2020

This work was completed under a memorandum of understanding (MOU) between the Government Data Analytics Center (GDAC) within the Department of Information Technology (NC DIT) and the University of North Carolina at Chapel Hill (UNC-CH) regarding the provision of services to support a study evaluating how to modernize and increase interoperability of the North Carolina Longitudinal Data System.

Table of Contents

Executive Summary

Background

North Carolina is at the forefront of a national movement to leverage longitudinal data systems (LDS) to inform policy decisions. Through the prior efforts of stakeholders—NC Department of Commerce (NCDOC), NC Community College System (NCCCS), NC Department of Health and Human Services (NCDHHS), NC Independent Colleges and Universities (NCICU), NC Department of Public Instruction (NCDPI), and the UNC System—much progress has been made in the collection of data and the creation of targeted data systems.

As the policy and program landscape becomes increasingly complex, there is an increased need for better data and analytics to help navigate complex decisions. An LDS can help North Carolina move from collecting data for compliance and accountability purposes to using data to support evidence-based policymaking, continuous improvement, and performance management. It can increase our knowledge of the opportunities and challenges that North Carolinians experience as they transition from early childhood, through the education system, and into the labor market. A North Carolina Longitudinal Data System—NCLDS—can enable the systematic use of evidence to guide decision-making and to improve citizens' lives.

LDS Milestones in North Carolina

1992 The Common Follow-up System (CFS) was initiated to provide information on the educational and employment outcomes of participants in publicly supported educational, employment, and training programs.

2011 The Early Childhood Integrated Data System (ECIDS) was initiated to track program participation and to inform policies and practices that produce better outcomes for children and families.

2012 North Carolina SchoolWorks (NCSW) was initiated to select and provide access to information on cohorts of students, schools, and program data over time.

2016 The NC Department of Information Technology's (NCDIT) Government Data Analytics Center (GDAC) was charged with developing an implementation plan to phase in the establishment and operation of a North Carolina Longitudinal Data System.

2018 Governor Roy Cooper requested that the NC Education Cabinet convene a working group to plan for a broader North Carolina Longitudinal Data System.

The Charge from the Education Cabinet

In 2019, GDAC entered into an MOU with the University of North Carolina at Chapel Hill (UNC-CH) to develop a strategic plan to modernize the NCLDS, grounded in the needs and priorities of policymakers, program leaders, and participating agencies. The team conducted over 40 interviews with stakeholders, including executive leadership, program managers, and technical leads within key agencies to understand their vision for an integrated state LDS, and the opportunities and challenges it would present.[1] The team also interviewed longitudinal data users and managers from other states to learn about their systems and processes. The perspectives of NC stakeholders, experts from other states, and evidence-based, best practices shaped this report and roadmap for building an NCLDS.

A Note on Terminology

Interviews with stakeholders revealed confusion and/or conflicting perceptions around the term "educational longitudinal data system" or ELDS. Some stakeholders used ELDS to refer to an individual existing LDS and other stakeholders used it to mean a broader, integrated longitudinal data system that would encompass the functionality of the three existing systems: CFS, ECIDS, and NCSW. The report will use the term **North Carolina Longitudinal Data System (NCLDS)** to identify a modernized longitudinal system that will build upon the three existing systems in North Carolina.

Shared Vision for NCLDS

The UNC-CH team interviewed stakeholders to understand their long-term goals and business priorities for a longitudinal data system. Stakeholders agreed that a comprehensive longitudinal data system would be a valuable tool for collecting and disseminating data to inform policy and program decisions.

Stakeholders shared a range of perspectives on data system specifics, but there was substantial consensus on the following points:

- NC's most pressing questions should drive the content and functionality of a modernized NCLDS to support data-informed decision-making;
- NCLDS partners should collaborate to define procedures and safeguards to protect individuals' privacy; to ensure data security; and to establish a transparent process for authorizing tiered access for different user groups;
- NCLDS should be designed to alleviate rather than to increase the workload for existing technical and program staff; and
- Analytical capacity should be strengthened to ensure that NCLDS data are transformed into actionable insights.

[1] We would like to thank the many people who generously shared their time, expertise, and perspectives during the research and writing of this report. For a full list of project interviews, see Appendices 1 and 2.

Trust as the Key Building Block

NC stakeholders viewed trust as the key building block to an effective NCLDS. Each acknowledged that existing agency "silos" undermine trust and limit awareness, understanding, and use of data across entities.

Stakeholders shared a range of perspectives about how trust factors into NCLDS modernization efforts. Trust bolsters relationships between individuals—from system leaders to program and data experts—to improve collaboration across agencies. Partners described how ambiguity around longitudinal data system leadership and vision have strained trust between the partners and have contributed to overall project fatigue.

Trust can be strengthened by a transparent, formalized process that authorizes access to and use of NCLDS data, and by a secure technical infrastructure, housed by a neutral entity, that safeguards high-quality, confidential data. All partners agreed that NCLDS governance would play a pivotal role in building trust and reducing silos.

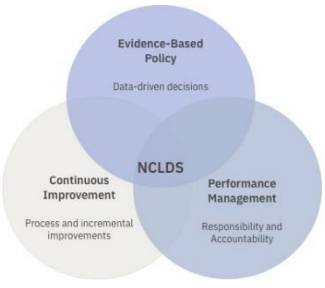
Recommendations

NCLDS should be designed to support evidence-based policymaking, continuous improvement, and performance management. These recommendations reflect an assessment of steps that North Carolina must take to establish a system with the capacity to inform decisions regarding policy and programs along the early childhood, Kindergarten-Grade 12 (K-12), higher education, and workforce continuum. The recommendations are organized topically (rather than sequentially) into eight categories. Collectively, they put the people and processes in place needed to support NCLDS. These conclusions are based on interviews with NC stakeholders, discussions with system experts in other states, and a review of best practices from recognized experts such as the U.S. Department of Education's State Longitudinal Data System (SLDS) project, the Data Quality Campaign, and Actionable Intelligence for Social Policy.[2]

The recommendations are designed to nurture trust by structuring stakeholder engagement, governance, analytical and research capacity, and sustainable investments to provide transparency and to facilitate collaboration across partner agencies and entities. The recommendations marked with  are part of a comprehensive strategy to build trust.

<u>Recommendation 1</u>: Affirm the NCLDS vision as a "system of systems" that links data from across agencies and over time to support evidence-based policy, performance management, and continuous improvement:

Figure 1: NCLDS Vision



- Eliminate use of the term ELDS in outreach and education efforts to ensure clarity regarding the relationship between NCSW and NCLDS.

- Use the term NCLDS to identify a "system systems" that links data across early childhood, K-12, postsecondary education, and workforce. Beginning

[2] For a list of electronic reference materials, see Appendix 7. Figure 1 adapted from The Integrated Data System Approach: A Vehicle to More Effective and Efficient Data-Driven Solutions in Government.

immediately, employ the term NCLDS consistently in documentation, communication, and related online resources.

- Revisit General Statute 116E to address ambiguity regarding NCLDS definition, location, and governance. NC stakeholders identified the following examples of ambiguity:
  - §116E is titled "Education Longitudinal Data System," although §116E-5 creates the "North Carolina Longitudinal Data System."
  - §116E-4 specifies the powers and duties of GDAC, but §116E-5(a) states that the system will be located administratively within the Department of Public Instruction.
  - §116E-4(a) directs GDAC to establish a committee on data quality but does not confer authority to establish a governance structure to coordinate participation of partner agencies.

Recommendation 2: Craft a stakeholder engagement strategy to build support for NCLDS:

Leverage Education Cabinet support for NCLDS to secure buy-in from the General Assembly and key system and agency decision-makers.

Engage a broad group of NCLDS stakeholders using an equity approach—policymakers, agency leaders, program managers, data contributors, legislative and fiscal analysts, postsecondary leaders, school districts, parents/families, and community leaders—to participate in NCLDS design and to identify questions that it should be able to answer.[3]

- Encourage contributing partners to promote NCLDS' potential to support data- and knowledge-driven solutions that address policy challenges along the early childhood, K-12, postsecondary education, and workforce continuum.
- Offer user-friendly information sessions for different user groups on NCLDS functionality, including dashboards, analytical tools, and other system outputs.

Recommendation 3: Establish NCLDS governance and organizational structures:[4]

Establish an Executive Board (by statute) comprised of senior leaders or their designees from each contributing agency. In the interim, appoint an Executive Committee (by administrative action) with the same representation.

[3] The Hunt Institute's Informed Decision-Making Collaborative, funded by the Bill and Melinda Gates Foundation, has convened North Carolina stakeholders to identify research questions and begin development of a shared research agenda.

[4] This will require clarifying the relationship between the NCLDS governance structure and existing LDS.

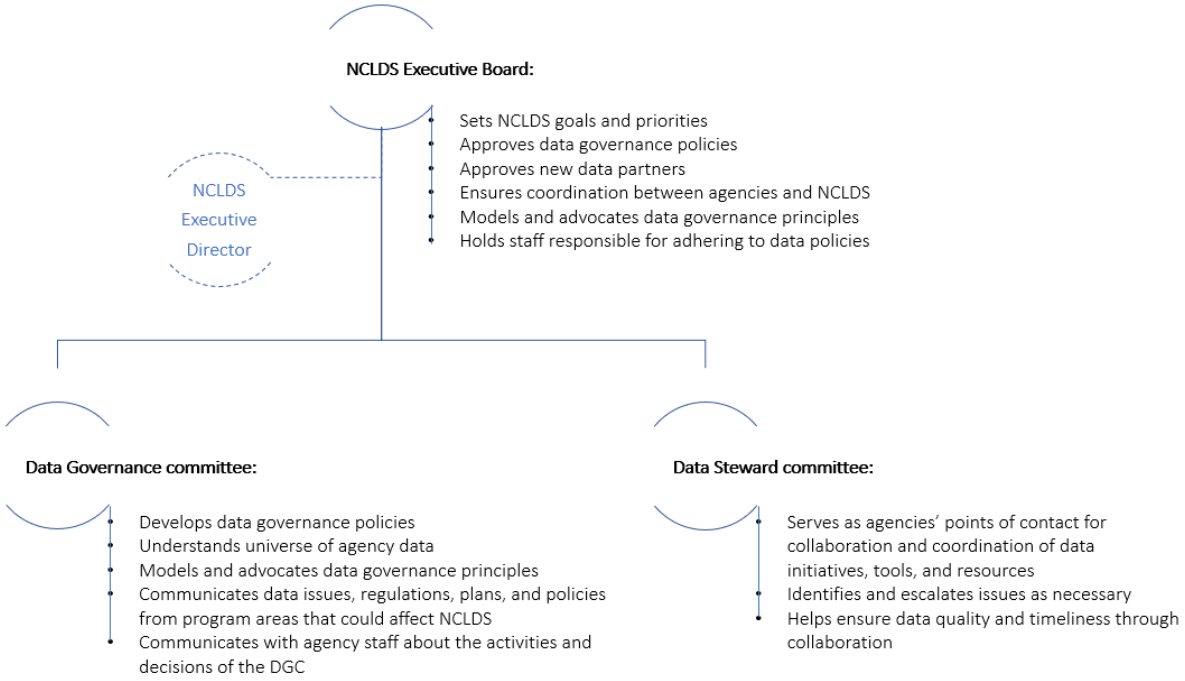
- Hire an NCLDS Executive Director to oversee the implementation and operation of the system and to act as a liaison between the Executive Board/Committee and NCLDS data governance committees.

- Name representatives from data-contributing agencies and entities to Data Governance and Data Steward committees.

- Adopt data governance policies that provide data stewards with control over the use of their data and that build trust in NCLDS as a partner in safeguarding that data.

- Designate GDAC as the administrative home for NCLDS, leveraging its standing as a neutral entity that has secured stakeholder trust and the trust of the NC General Assembly. This role would be consistent with GDAC's responsibilities to support NC's efforts to improve longitudinal data systems.[5]

Figure 2: Proposed NCLDS governance structure



[5] As identified in statute, GDAC's role in statewide data integration and sharing is to identify data integration and business intelligence opportunities that improve the efficiency and effectiveness of state agencies, departments, and institutions (G.S. 143B-1385). GDAC currently hosts ECIDS, NCSW, and CFS. Chapter 116E grants GDAC enumerated powers and duties with respect to operation and oversight of North Carolina's Longitudinal Data System.

Recommendation 4: Prioritize NCLDS requirements for a phased implementation:

- Prioritize NCLDS users, starting with policymakers, agency and program leaders next, and external researchers next.

- Document and prioritize functional requirements for a flexible system of inquiry, including standard and customizable dashboards, reports, data stories, knowledge visualizations, research-ready data sets, metrics, and analytics. Identify "low-hanging fruit" (e.g., research-ready data sets) for initial implementation.

- Identify the current data contributors to ECIDS, NCSW, and CFS as initial data contributors for NCLDS.

- Identify and prioritize additional data sources for inclusion in NCLDS over time (e.g., National Student Clearinghouse, U.S. Census Bureau, U.S. Bureau of Labor Statistics data, NC Families Accessing Services through Technology (NC FAST), Juvenile Justice, or NC licensing board data, see Appendix 4).

Communicate the plan to NCLDS stakeholders.

Recommendation 5: Develop a system architecture to meet NCLDS functional and technical requirements:

- Design a scalable NCLDS architecture and system infrastructure to meet the functional priorities identified by NCLDS governance.

Include NCLDS stakeholders in the system design process to ensure that the technical infrastructure has the flexibility to support NCLDS' short- and long-term goals and the safeguards to address concerns about data quality, data security, and data privacy.

Form a working group to broaden and to formalize discussions about data quality issues that need to be resolved.

Coordinate and further develop data privacy practices and procedures with data contributors.

- Establish a centralized system for entity resolution. In the interim, NCLDS can use source systems' unique identifiers and manual crosswalks. Over the longer term, leverage GDAC's Enterprise Entity Resolution (EER) to provide a consistent and scalable mechanism for linking data and adding data sources over time.[6]

[6] Research and development on EER methodology are already underway.

Recommendation 6: Build analytic capacity to support NC's evaluation and research priorities:

Hire a Director of Analytics and Research early in the process, who will report to the Executive Director, to oversee the internal analytics team, as well as to serve as the interface with external research partnerships. The Director of Analytics and Research would implement a plan based on the priorities set by the NCLDS Executive Board/Committee.

Establish an internal analytics team with subject matter expertise to support evidence-based policymaking, continuous improvement, and performance management.

Implement a staffing strategy that offers training rotations and/or dual employment in NCLDS and contributing agencies and entities to build cross-agency program and data fluency for new and existing analytic staff.

Recommendation 7: Develop an external research agenda to prioritize data requests and the formation of research-practice partnerships:

Create a collaborative, NC (internal) policy research working group composed of agency representatives, subject matter experts, and practitioners to develop a research agenda[7] and to enhance understanding of cross-agency priority areas.

- Design a framework that prioritizes data requests that are aligned with the research agenda, and that supports the formation of (external) research-practice partnerships.

Establish research approval requirements to ensure that (a) originating agencies approve data use and provide aggregated or de-identified data, (b) researchers maintain active IRB approval status and comply with NCLDS research review requirements, and (c) research findings are made available to the public as deemed appropriate by NCLDS governance committees.

- Create research-ready datasets to guide external researchers and applied policy analysts toward priority topics within the research agenda, providing researchers with high-quality data while making more efficient use of staff time than responding to "one-off" requests.

- Form research-practice partnerships to augment the capacity of the state and its agencies to undertake large-scale research and evaluation initiatives.

[7] This group would continue the work started under the Hunt Institute's Informed Decision-Making Collaborative.

Recommendation 8: Develop an action plan to identify immediate next steps and to secure sustainable funding for NCLDS design, implementation, and operation:

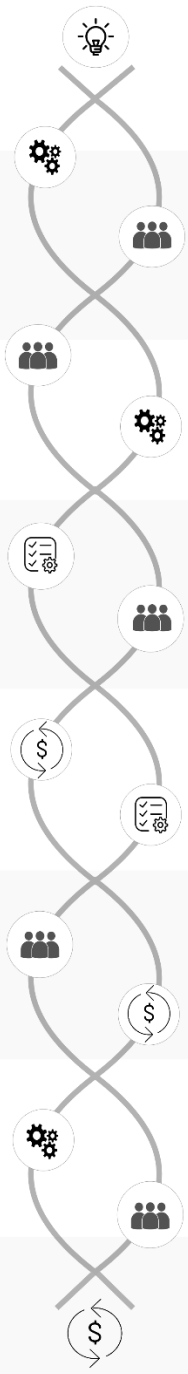
- Move NCLDS forward in the short-term by: (a) asking the Education Cabinet to endorse an action plan, (b) reallocating existing resources to staff key NCLDS positions in the interim, including an Executive Director, (c) establishing the Executive Committee (administratively) and the Data Governance and Data Steward committees, (d) prioritizing functional requirements and identifying "low-hanging fruit" for initial implementation, and (e) estimating the technical and non-technical resources needed from each agency.

- Advocate for state appropriations with support from partner agencies to fund: (a) NCLDS design, (b) initial NCLDS implementation and source system upgrades, (c) system maintenance and operations, and (d) staffing costs for positions dedicated to NCLDS analytic and technical operations.

- Leverage federal and private foundation grant opportunities to support key priorities of stakeholders in piloting new functionality, adding data sources, and forming research-practice partnerships.

A Roadmap to NCLDS

Putting people and processes in place



Affirm the NCLDS vision as a "system of systems" that links data from across agencies and over time to support evidence-based policy, performance management, and continuous improvement

Craft a stakeholder engagement strategy to build support for NCLDS

Establish NCLDS governance and organizational structures

Prioritize NCLDS requirements for phased implementation

Develop a system architecture to meet NCLDS technical and functional requirements

Build analytic capacity to support NC's evaluation and research priorities

Develop an external research agenda to prioritize data requests and the formation of research-practice partnerships

Develop an action plan to identify immediate next steps and to secure sustainable funding for NCLDS design, implementation, and operation

Section 1: The Past, Present, and Future of Longitudinal Data Systems in North Carolina

Establishing a Statewide Longitudinal Data System

North Carolina is at the forefront of a national movement to leverage longitudinal data systems to inform policy decisions. NC's efforts date back to the 1990s, when a research shop, in what is now the Department of Commerce (NCDOC), led a multi-agency collaborative to build the Common Follow-up System (CFS), one of the nation's first education and workforce data repositories. In 2013 and again in 2019, the DOC received federal grants from the Workforce Data Quality Initiative to improve and to expand CFS. Over the last two decades, other NC agencies have secured federal funding to develop targeted longitudinal systems. In 2007 and again in 2012, the NCDPI secured competitive grants to build a PK-13 system (the Common Education Data Analysis and Reporting System, or CEDARS) and a P-20W system (NC SchoolWorks, or NCSW). In 2011 and again in 2020, the NCDHHS received funding for early childhood integrated data systems (ECIDS). Individually, these efforts set the standard for early childhood, education, and workforce data systems.

In 2018, Governor Roy Cooper reconvened the Education Cabinet and challenged participating members to increase data sharing for decision-making. This challenge generated interest in and momentum for connecting the state's existing longitudinal data systems—a "system of systems" that would address the goals specified under General Statute 116E[8] and enable NC to make data-driven decisions across the early childhood to education to workforce continuum.

To achieve this goal, NC needs a roadmap grounded in the needs and priorities of policymakers, program leaders, and participating agencies. A University of North Carolina at Chapel Hill team interviewed dozens of NC stakeholders to understand their vision and expectations for an integrated statewide longitudinal data system and the opportunities and challenges it would present.[9] The team also interviewed longitudinal data users and managers from other states and documented their systems and processes. Stakeholder perspectives, interviews with other states, and best practices from national experts shaped this report and recommendations for a way forward.

Origins and Evolution of NC's (Existing) Longitudinal Data Systems

The path forward must leverage and build upon existing efforts, so we begin with an overview of NC's three existing longitudinal systems: CFS, NCSW, and ECIDS. These systems originated at different points in time: each with varying degrees of cross-agency participation and coordination; and each designed to

[8] As specified in statute, that system must have the capability to:

- Facilitate and enable the exchange of student data among agencies and institutions;
- Generate timely and accurate information about student performance that can be used to improve education systems and guide decisions makers at all levels;
- Facilitate and enable linkage of student data and workforce data;
- Serve as a data broker for education and workforce data.

[9] We would like to thank the many people who generously shared their time, expertise, and perspectives during the research and writing of this report. For a full list of project interviews, see Appendices 1 and 2.

meet specific agency needs and/or federal grant requirements. As NC moves forward, these systems will form the foundation of an integrated statewide longitudinal data system.

Common Follow-Up System (CFS)

The CFS was created in 1992 as a cooperative venture among participating state agencies to provide information on the educational and employment outcomes of participants in publicly supported educational, employment, and training programs.[10] Over the first few years of operation, the system was converted from a single-year matching system to a longitudinal database.

In 1995, the NC General Assembly (NCGA) enacted legislation that established CFS by law. The statute defined system participation; established and assigned operational and evaluative responsibilities; mandated data integrity and confidentiality; and outlined reporting requirements and schedules. The former Employment Security Commission (ESC) was delegated operational (and, later, evaluation) responsibility while the Office of State Budget and Management (OSBM) was charged with analysis.

In 2011, the NCGA enacted legislation that transferred the ESC to the Department of Commerce. The NCDOC's Labor and Economic Analysis Division (LEAD) took on responsibility for CFS, including the associated analytics.

In 2012, the NCGA enacted reforms to the state's workforce development system, directing Commerce to strengthen the CFS and collaborate with the Commission on Workforce Development to develop performance measures for workforce development using information from CFS.

In 2014, the NCGA required Commerce to develop a plan for the transfer of the information and capabilities of CFS to the NCDIT's GDAC. Since 2015, LEAD has worked with GDAC and contributing agencies to carry out data processing, submission, validation, and approval in a secure GDAC portal. LEAD will continue to partner with GDAC to improve the quality of data matching capabilities through the Enterprise Entity Resolution process, which is still under development and will be used for data matching where common keys for individuals are not available, and to expand visual analytics for contributors' data reporting. CFS is a warehouse system, and approved data are loaded to individual contributor warehouses in the GDAC environment.

CFS Reporting and Tools

In addition to the annual CFS Operational Report and the biennial CFS Evaluation Report, data in the CFS has been utilized to support several reporting tools and dashboards. NC's Tool for Online Workforce and Education Reporting (NC TOWER) is a public-facing, web-based system that uses CFS data to report employment and wage outcomes for graduates from the University of North Carolina System schools and the North Carolina Community College System. LEAD also uses CFS data to feed the Labor Supply/Demand Analyzer (which shows the alignment of the state's higher education

[10] Participating agencies include the Department of Commerce, the UNC System, NCCCS, NCDPI, the NC Department of Public Safety (NCDPS), and NCDHHS. For additional details, see A Report on the Operations of the North Carolina Common Follow-up System.

system to the needs of the labor market) and the Workforce Development Board's dashboard (which provides outcomes and economic impact of individuals served by the local career centers). In addition, CFS data are utilized to create a set of performance measures to assess the state's workforce development system programs for the NCWorks Commission. Currently, various agencies partner with LEAD to help evaluate the effectiveness of their programs, using data that they have provided to CFS.[11]

North Carolina SchoolWorks (NCSW)[12]

NC SchoolWorks is a federated longitudinal data system designed to select and provide access to information on cohorts of students, schools, and program data over time. NCSW is the most recent NCDPI initiative to improve the state's management of student-level data.

In 2007, NCDPI received a $6 million federal grant from the U.S. Department of Education's Statewide Longitudinal Data Systems Grant Program. Under this grant, NCDPI developed the Common Education Data Analysis and Reporting System (CEDARS), which included the creation of a Unique Statewide Identifier (UID) for both students and teachers. CEDARS is NCDPI's PK-13 data warehouse.

In 2012, NCDPI was awarded an additional $3.6 million federal grant to develop a federated, P-20W statewide longitudinal data system, formalized in statute to include "individual-level student and workforce data from all levels of education and the State's workforce." To date, federal funds have been used (a) to establish a data broker, (b) to develop a distributed query system to facilitate data sharing among NCSW partners, (c) to make necessary modifications to source systems to support UID matching, and (d) to adopt memorandums of understanding among the partners.

At this time, NCSW can provide information from NCDPI, the UNC System, NCCCS, and CFS. NCSW cannot produce data from NCDHHS (via ECIDS)[13] or the North Carolina Independent Colleges and Universities.

NCDPI is the business owner of NCSW. GDAC hosts the application and provides technical support. NCDPI and GDAC are working together to coordinate ongoing NCSW development.

NCSW Reporting

NCSW was designed for researchers, and it includes an automated data request process, but in its current state, NCSW does not include reporting tools.

[11] CFS is working on a pilot with three NCICU member institutions to look at employment patterns for graduates.
[12] Participating agencies include NCDPI, the UNC System, NCCCS, NCICU, NCDHHS, and the Department of Commerce.
[13] The current ECIDS transition to a SAS platform should resolve some of the technical barriers that have prevented inclusion of early childhood data. After the transition, some additional development will be required to complete the connection.

Early Childhood Integrated Data System (ECIDS)[14]

ECIDS includes data on early childhood education, health, and social services. ECIDS was developed with the support of a $6.9 million federal grant in 2011. ECIDS integrates data from programs that serve very young children—many starting at birth—to track program participation and inform policies and practices that produce better outcomes for children and families.

ECIDS was first developed as a federated system, using the UIDs to enable integration with K-12 data. Currently, ECIDS is transitioning to a new data warehouse, which, along with select standardized reports, is scheduled to launch in 2020.

ECIDS Reporting

ECIDS data are available through an interactive web portal that provides aggregate statewide reports and that can be customized by county, child demographics, and state fiscal year. ECIDS has a dedicated, private data request portal for research data requests.

Toward a North Carolina Longitudinal Data System

CFS, NCSW and ECIDS were developed at different points in time, by overlapping coalitions of partner agencies, responding to specific federal funding opportunities. All three support data-informed decision-making to improve services and outcomes for North Carolinians. However, these efforts do so with data systems that reflect the grant criteria and priorities of participating federal agencies and the nuances of the populations and programs they serve, resulting in data sharing and connectivity issues.

In 2012, the NCGA established a P20W system, and laid the groundwork for a more coordinated approach in General Statute 116E. In a 2016 revision, the NCGA charged GDAC (which currently maintains the technical infrastructure for CFS, NCSW, and ECIDS) with specific responsibilities for a statewide longitudinal data system.

In 2018, the reconvened Education Cabinet assembled a cross-agency board to promote data sharing. This led the Education Cabinet to create a Steering Committee to work with GDAC on developing an integrated statewide longitudinal data system. The Steering Committee supported efforts to: (a) draft the rules and agreements necessary to facilitate data sharing; and (b) engage stakeholders in defining the need and vision for a statewide longitudinal data system. To help define the need and vision, GDAC entered into an MOU with UNC-CH to conduct a modernization study in collaboration with the Steering Committee.[15]

Building the Roadmap: Methods

UNC-CH was charged with developing a roadmap responsive to the needs and priorities of policymakers and participating agencies. The goal, as specified by the MOU, was to determine how to enhance the availability and access of cross-agency and sector-actionable intelligence, data, and metrics to support

[14] Participating agencies include NCDHHS' Division of Child Development and Early Education, Division of Public Health, and Division of Social Services; and NCDPI's Office of Early Learning and Head Start/Early Head Start.
[15] The Education Policy Initiative at Carolina (EPIC) is an initiative within the Department of Public Policy at the University of North Carolina at Chapel Hill.

data- and knowledge-driven solutions that address complex social problems along the early childhood, education, and workforce continuum.

The UNC-CH team started by clarifying objectives and priorities with the Steering Committee at a kickoff meeting in October 2019. They conducted interviews with GDAC and reviewed public documents to understand better the functionality, products, and stakeholders of NC's existing longitudinal data systems.

The UNC-CH team developed an interview plan and tailored protocols for soliciting input from a broad range of stakeholders in the state. They conducted a series of interviews with leaders and technical and programmatic teams from participating agencies, as well as staff from other state and nonstate entities, to explore the business needs for an integrated longitudinal data system. They requested input on the desired functionality and key user groups, and perceptions about the opportunities and challenges presented by a statewide longitudinal data system. They presented briefings from these interviews at regular meetings with GDAC and the Steering Committee throughout the fall and winter of 2019-2020.

Early in the project, the UNC-CH team surveyed best practices and other states' longitudinal data systems, consulting the Steering Committee to select a handful for deeper analysis. During sessions with NC stakeholders, interviewees expressed interest in how other states developed a vision; engaged stakeholders; designed governance structures; managed research agendas, data requests, and system usage; and planned for sustainability. These questions shaped the interview protocol for comparison states. The UNC-CH team interviewed a total of eight states—Connecticut, Georgia, Kentucky, Maryland, Minnesota, Rhode Island, Washington State, and Wisconsin—though the focus and specificity of those interviews varied in accordance with the tenure and expertise of the interviewees.

The UNC-CH team synthesized themes from more than 40 interview sessions to highlight key takeaways. Input from stakeholders, findings from other states, and published reports on best practices drove the outline for this report and helped to position the guideposts for a statewide longitudinal data system roadmap.[16]

[16] A draft of this report was distributed to the Steering Committee for feedback prior to publication.

Section 2. Creating a Vision and Mission Out of Acronyms

What is a Longitudinal Data System (LDS)?

An LDS is a data system that links individual-level data across sectors and over time to inform state policy and practice around transition points. An LDS can enable performance management, continuous improvement, and evidence-based policymaking; broadly, it is the systematic use of evidence to guide decision-making in government to improve citizens' lives.

To support those functions, LDS outputs can be disseminated to different users via carefully governed processes and procedures, for example:

- Public-facing data dashboards;
- Standard reports that are regularly updated and posted online;
- Interfaces that enable users to submit customized queries for aggregate-level data; and
- Portals that allow users to request access to de-identified, individual-level data to drive analytics and research regarding key policy questions.

The LDS landscape varies from state to state because each LDS reflects its state's policy priorities and variations in fiscal and technical capacity.

Clarifying Ambiguity around ELDS

The UNC-CH team met with stakeholders to understand their long-term goals and business priorities for a longitudinal data system. However, these interviews reflected substantial confusion around the term "educational longitudinal data system" (ELDS) that was introduced in 2012 in General Statute 116E.

- A majority of interviewees associate ELDS with one of the existing longitudinal data systems—most often NCSW.
- A minority of interviewees interpreted ELDS as a broader, integrated longitudinal data system that would encompass the functionality of the three existing systems.

When the UNC-CH team asked questions about governance, some stakeholders referred to NCSW governance, while others referred to governance of a broader longitudinal system. When the UNC-CH team asked about future capabilities, some stakeholders responded within the confines of existing NCSW functionality, while others addressed the opportunities presented by an integrated system.

A coordinated effort to modernize NC's existing longitudinal systems will require that all stakeholders adopt a shared vision and standard terminology. Our recommendations will address this issue directly because on-going ambiguity prevents partner agencies from developing the common vision and trust needed to move a data system forward.

To minimize confusion, this report will refer to existing systems by their respective names—CFS, ECIDS or NCSW. The report will use the term **North Carolina Longitudinal Data System (NCLDS)** to identify a future longitudinal system that will build upon the three existing systems in North Carolina.

Creating a Shared Vision for NCLDS

Across all agencies, stakeholders agreed that a comprehensive NCLDS would be a valuable tool for collecting and disseminating data to inform policy and program decisions.

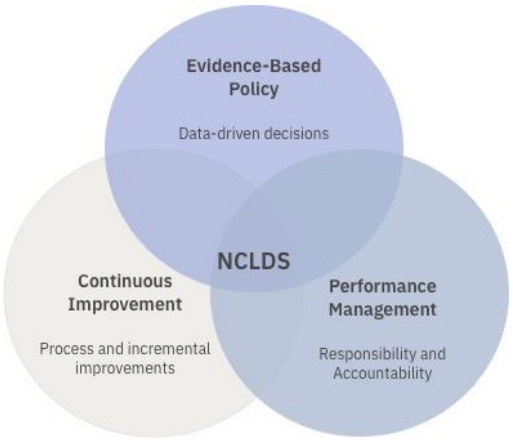
Stakeholders shared a range of views on system specifics, but there was substantial agreement that:

- NC's most pressing policy questions should drive the content and functionality of a modernized NCLDS to support data-informed decision-making;
- Agencies should collaborate to define procedures and safeguards to protect individuals' privacy, ensure data security, and establish a transparent process for authorizing tiered access for different user groups;
- NCLDS should be designed to alleviate rather than to increase the workload for existing technical and program staff; and
- Analytical capacity should be expanded to ensure that NCLDS data are transformed into actionable insights.

As envisioned by the dozens of stakeholders interviewed, NCLDS would provide high-quality, aggregated and record-level data:[17]

- To help answer some of the state's most complex policy questions around (a) the transitions, trajectories, and outcomes of young children, students, and workforce participants and (b) the impact of system shocks such as the COVID-19 pandemic and economic downturns;
- To support performance management and continuous improvement of programs designed to improve early childhood, education, and workforce outcomes; and
- To equip policymakers, other state leaders, and agencies and entities with information to plan strategically on the future needs of North Carolina's young children, students, workforce, and industries.

Figure 3: NCLDS Vision



Stakeholder perspectives chart a roadmap for an NCLDS that will be designed to support a range of uses by entities involved in these functions.

[17] Graphic adapted from The Integrated Data System Approach: A Vehicle to More Effective and Efficient Data-Driven Solutions in Government.

Prioritizing NCLDS Uses

NCLDS can be designed as an effective and efficient vehicle to analyze iteratively and to inform evidence-based policy, continuous improvement, and performance management. Identifying and prioritizing NCLDS uses and use cases is critical to determining system requirements and to creating a phased implementation plan that includes the requisite data elements, reporting and querying functionality, and user support. NC stakeholders identified three key uses for NCLDS and prioritized them (from highest to lowest): policy development, program management, and research.

Policy Development

Overwhelmingly, stakeholders report that, above all else, NCLDS should serve entities involved in developing evidence-based policies. Those entities may include state legislators, agency leaders, governing boards, interagency commissions, and public-private partnerships (such as the B-3 Council, myFutureNC, and the NCWorks Commission) that support the policymaking process. With easier and increased access to comprehensive, integrated data and sophisticated analytics, these users could make data-informed decisions about policies, goals, and resource allocations.

Program Management

According to NC stakeholders, the second NCLDS use centers around program management and decision support. Entities involved in programs' continuous improvement and performance management efforts may include agency and division leaders, program leads, local educational agency (LEA) and community college leaders, and others who guide program choices to achieve policy goals. Although this group already has access to existing operating systems,[18] NCLDS could provide efficient access to key data and reports, with better information about how program inputs and outputs connect with related programs, services, and supports.

Research

NC stakeholders had divergent perspectives about providing access to NCLDS data for external academic and policy research. Some interviewees view academic and applied policy researchers as critical partners in providing expertise and capacity that are in short supply in state government. Others view researchers as a burden on existing resources, with limited understanding of program details and a tendency to misinterpret data in the absence of sufficient guidance from program experts.

Most agreed that an NCLDS should include elements that facilitate productive partnerships between academic and applied policy researchers and government entities, for example:

- Establishing a research agenda, based on input from NC policymakers and agency leaders, that specifies topics for which they seek actionable analysis to inform policy and program decisions that directly benefit NC citizens;

- Providing research-ready data sets that are aligned with that agenda;

[18] A few stakeholders mentioned the potential for NCLDS to be used for regulatory reporting, but most did not see regulatory reporting as a core NCLDS function and thought that source systems would continue to perform that function.

- Streamlining request processes and prioritizing access for proposals that are related to the research agenda and priorities; and

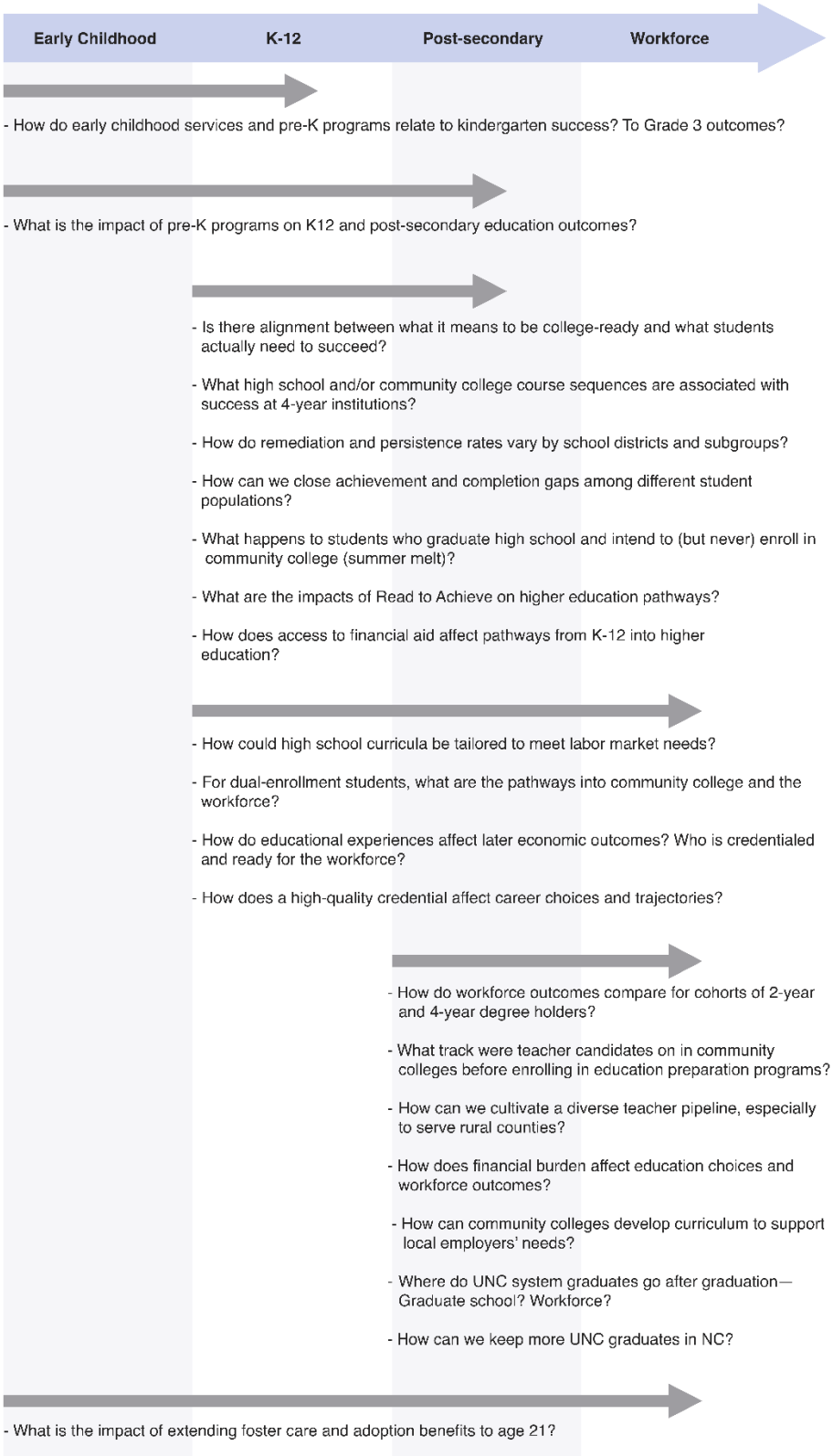
- Publishing standard data definitions, reference materials, and guidance on data use.

Defining NCLDS Use Cases

Stakeholders described the ways that their agencies and organizations would use NCLDS to inform policy and program decisions. Most of these use cases centered around the need for better data about transition points on the early childhood, education, higher education, and workforce continuum and related outcomes. The use cases highlight some of the limitations of existing systems.

The list of use cases in Figure 4 is illustrative, and it will expand as comprehensive data, enhanced analytics, and a changing policy landscape allow new applications for NCLDS to surface. Indeed, the COVID-19 pandemic, which surged during the writing of this report, produced a flood of potential use cases and underscored the need for a modernized NCLDS.

Figure 4: Example NCLDS Use Cases



Section 3. Building Blocks

NC stakeholders voiced strong support for a modernized LDS that links data across agencies and over time to inform better state policy and program decisions. In interviews, stakeholders talked about elements critical to the success of NCLDS. These "building blocks" align with published best practices and evidence from interviews with other states.

Shared Vision and Mission

NC stakeholders emphasized that a successful NCLDS begins with a clearly articulated vision about its purpose; having a clear mission and vision statement that underscores the rationale for building a system is fundamental to the project's overall success. That vision lays the groundwork for securing the support of state legislators and agency leaders, building trust among partners, and making decisions related to the function and structure of the system.

Executive and Legislative "Buy-In"

NC stakeholders underscored the importance of building support for NCLDS among state leaders—support that spans the Office of the Governor, the General Assembly, the Education Cabinet, state agencies and other key stakeholders. To deliver on the promise of data-informed decision-making, NCLDS will require a significant commitment of human and financial resources. NC agency leaders recognized the critical role they will play, within and across their organizations, securing support from the General Assembly and agency staff to prioritize resources for NCLDS.

Interviews with other states suggest that high-profile, executive support can help build consensus and muster resources around a shared vision. For example, Maryland and Washington State benefited from strong gubernatorial support to align resources and to accelerate progress on their longitudinal data systems. In interviews with these states, officials agreed that executive leadership could play an important role in building and sustaining momentum around a shared vision. Washington State also reported that legislative champions were the key to sustainability, a perspective echoed by Minnesota, where the legislature has funded SLEDS since 2014-2015. Wisconsin officials shared that they spend a lot of time demonstrating the value of their system to ensure that buy-in can endure political change.

Trust among NCLDS Partners[19]

NC stakeholders viewed trust among partners as the cornerstone of an effective NCLDS. They acknowledged that existing agency silos undermine trust and limit awareness, understanding, and the use of data across entities. They described how ambiguity around LDS leadership and vision have strained trust between the partners and contributed to enduring project fatigue.

Stakeholders shared a range of perspectives about how trust factors in to NCLDS modernization efforts. Trust bolsters relationships between individuals—from system leaders to program and data experts—to improve collaboration across agencies. Trust can be strengthened by a transparent, formalized process that authorizes access to and use of NCLDS data, and by a secure, technical infrastructure, that is housed by a neutral entity, and that safeguards high-quality, confidential data.

[19] P-20W+ Data Governance: Tips from the States

Most stakeholders agreed that the technical infrastructure, although complex in its own right, would be simpler to build than trust. They thought that NCLDS governance structures would play a pivotal role in building trust and reducing silos.

Interviews with officials from other states offer evidence on how trust factored into their experiences. For example, a Rhode Island official identified lack of trust as the state's most significant implementation challenge. In Washington State, officials noted that statute mandated the agencies' participation, which they felt sidestepped some trust-related roadblocks to collaboration.

Trust in NCLDS Data

Trust among NCLDS partners paves the way for productive conversations and needed agreement on a range of issues—including data quality, security, privacy, use, and interpretation—that collectively build trust in NCLDS data.

Data Quality[20]

Data quality is a multidimensional measurement of the adequacy of a variable or dataset, including accuracy, completeness, consistency, and timeliness, which can be affected by multiple factors in the assessment, collection, and analysis processes.[21]

As such, data quality is a core building block of an effective longitudinal data system. Issues such as missing values and inconsistent naming conventions can impact the reliability of data within a dataset. Furthermore, variations in quality across datasets can inhibit successful linking of key data elements needed for NCLDS analyses. Many stakeholders shared additional data sources that may be beneficial to add to NCLDS (see Appendix 4), but the first step is to ensure the quality of the data at its source. Stakeholders also reported that a pre-existing constraint on an LDS may be the quality of historical data that predates quality control processes implemented over the last decade.

Stakeholders reported that the way that a data element is used within an agency could drive the quality of the data element. For example, NCDPI assigns a UID to all K-12 students, which it uses to link records across datasets and over time. Given the importance of this variable across its subsystems for regulatory reporting, NCDPI dedicates substantial time and resources to ensure fidelity across cases and over time. In contrast, the NCCCS and the UNC System use the UID only for matching data in NCSW and can tolerate a relatively higher rate of missing or invalid UIDs.

To assess data quality, NCDPI recently conducted a quality-matching test using UIDs, in which bulk data were uploaded from institutions of higher education (IHEs). The entity resolution results showed that 400,000 out of 8 million records (5%) did not match for UNC system schools, and 200,000 out of 6 million records (3%) did not match for community colleges.

Other stakeholders conveyed the challenge that students enrolled in multiple IHEs might end up with conflicting data and/or data that cannot be linked. Without reliable UID matching, it is difficult for IHEs to follow students who begin their postsecondary work at a community college and transfer into the UNC system. Stakeholders also talked about the challenges of entity resolution in existing LDSs when it comes to records that are near matches. Some stakeholders identified this as a significant burden on

[20] Data Quality: Striking a Balance
[21] Data Use Standards: Professional Behaviors

agency resources and advocated for moving the responsibility out of the agencies. Standards will need to be developed, adopted, and enforced to ensure a consistent quality of data elements and a process will need to be adopted for NCLDS entity resolution.

Based on these examples, NCLDS governance will need to agree on data quality standards that will apply across agencies. Enterprise Entity Resolution can assist in this effort, although NCLDS governance should explore the use of other matching mechanisms to support data integration.

Data Security[22][23]

Data security is defined as the protection of data from unauthorized (accidental or intentional) modification, destruction, or disclosure.[24]

Stakeholders affirmed that a future NCLDS must establish data security requirements to protect individual privacy and to maintain public trust, but there was no consensus on the specifics of how the requirements would be enforced. Many interviewees recommended providing training on security protocols because common standards around data security do not yet exist across agencies and entities. Stakeholders suggested that NDAs and MOUs would provide additional reassurance through legal protection around data security and use.

Stakeholders also shared concerns about legal ramifications in the event of an NCLDS data breach. Currently, liability coverage in NC differs between public and nonpublic institutions. State agencies are covered by sovereign immunity, meaning they can only be sued for certain events when performing work required by statute. Nonpublic entities that participate in NCLDS, including NCICU members, do not have sovereign immunity, but do receive some protection from liability under G.S. § 116-229.1, which provides that NCICU and private colleges and universities will not be liable for breaches of confidentiality caused by acts or omissions of State agencies and others who receive their data. However, nonpublic entities remain more vulnerable than public entities in the event of a legal action because they cannot rely on the state to provide legal representation. As a result, it is particularly important to nonpublic entities that NCLDS comply with all provisions of the law. NCICU interviewees report that required rules must be in place for their member institutions to participate in NCLDS to minimize their potential liabilities.

In the event that a legal action requires financial restitution, IT contractors are liable for damages up to two to three times the value of the contract.[25] Although NCICU members could incur significant legal expenses from legal action related to data breaches, these protections are considered important to member participation in NCSW or NCLDS.

[22] Traveling Through Time: The Forum Guide to Longitudinal Data Systems
[23] Best Practices for the Design and Implementation of Data Privacy and Security Programs; Safeguarding Data; Traveling Through Time: A Forum Guide to Longitudinal Data System
[24] SLDS Glossary
[25] G.S. 143B-1350(h1)

***Data Privacy*[26]**

The core elements of data privacy cover how an individual's personally identifiable information (PII) or other confidential personal information are stored, accessed, used, presented, and governed. This includes any and all information that can be used to identify, locate, or contact an individual.

Stakeholders emphasize that there must be trust among the partners, as well as from the public, that a NCLDS system can guarantee data privacy and confidentiality. This includes a commitment to ensuring that personally identifiable information is collected, maintained, used, and disseminated in a way that respects privacy, ensures confidentiality and security, and promotes access to data for policy development and implementation.[27]

Stakeholders report that privacy protections are necessary for a modernized NCLDS. Privacy concerns range from data collection to aggregating and reporting data. For example, interviewees note that these issues are particularly important as users become more sophisticated, and users may be able to triangulate information to ascertain an individual's identity.

An example of this concern was wage data collected in CFS.[28] Some wage data are considered trade secrets, and, with a few exceptions,[29] are federally restricted from being shared outside of government.[30] The confidential nature of information contained in CFS mandates the use of strict safeguards in the collection, storage, and use of the data. Stewards of wage data were concerned that NCLDS users may not all have the awareness or capacity to safeguard appropriately that data.

Stakeholders also expressed the need for a system that controls the level of access to data, prioritizing the protection of an individual's privacy as well as protecting the data stewards. Safeguarding gateways to data access is essential, but it can be resource intensive because it relates to human capital and system capacity.

CFS data are stored within the GDAC environment and access requires individual user data access profiles, as well as individual user IDs and passwords. At the time of system enrollment and with every data release, participants are informed of the confidential nature of the data and the legal restrictions on its use. Following the completion of the data loading process, CFS staff analyzes the data, and the resulting findings are utilized in the development and production of reports and other research products. The findings reported in these research products are subject to data suppression procedures that prevent the disclosure of personally identifying information.[31]

---

[26] How to Engage and Train Stakeholders Regarding Privacy and Security Best Practices; Identifying SLDS Users and Their Information Needs

[27] Data Stewardship Executive Policy Committee

[28] Sources and Linking Strategies for Employment Data

[29] Disclosure of unemployment compensation information is permissible in certain cases if authorized under state law and if the disclosure does not interfere with efficient administration of the state's unemployment compensation law. For additional information, see Legal Information Institute.

[30] NCICU expressed concerns that current restrictions prohibit nonstate entities like NCICU from accessing individual level wage data, thus limiting the data and analyses that NCICU could conduct using NCLDS.

[31] A Report on the Operations of the North Carolina Common Follow-up System

Interviewees report that confidence in data stewardship can be enhanced by transparent decision-making policies and processes that determine the type of data available to different entities (identified, de-identified, and aggregate), and the level of access made available (discrete subsets or access to full database). Stakeholders report that this can be a challenge, particularly with external researchers who want identified, individual-level data to track individuals over time, or when agencies, who may be granted access to aggregate-level data only, want to examine differences between subgroups. Anonymizing, blurring, sampling from larger data populations, and creating synthetic data sets could expand research uses while protecting privacy.

Based on our interviews, other states employed different decision-making policies and criteria to determine the level of access. For example, in Minnesota,[32] a complete list of individuals with access to their ELDS is maintained by the state IT agency and contributing state agencies. Access is approved by appropriate leadership as defined for each of six access levels, which are delineated based on the role of the data requester, ranging from read-write access for IT staff to anonymous reports produced for the public. For each level, approved staff must fulfill training requirements that are established by the state IT department and ELDS Executive Committee. Kentucky also uses a tiered system of access; however, levels are distinguished based on the purpose of data use, rather than on the role of the requestor.[33] This ranges from identified individual data used for the purpose of matching and linking records, aggregate data for measuring education and workforce programs, and aggregate data provided to the public through an externally facing website.

In addition to NC's specific rules for addressing privacy and liability issues, training for all users of an NCLDS about the Family Educational Rights and Privacy Act (FERPA),[34] the Health Insurance Portability and Accountability Act of 1996 (HIPAA),[35] UIDs, restrictions on distribution of wage data,[36] and the use of confidentiality agreements would provide additional assurances of compliance.

[32] Minnesota Statewide Longitudinal Data System: Data Access and Management Policy

[33] Kentucky Center for Education and Workforce Statistics: Data Access and Use Policy

[34] FERPA (20 U.S.C. § 1232g; 34 CFR Part 99) is a federal law that protects the privacy of student education records. The law applies to all schools that receive funds under an applicable program of the U.S. Department of Education. FERPA gives parents certain rights with respect to their children's education records. These rights transfer to the student when he or she reaches the age of 18 or attends a school beyond the high-school level. Students to whom the rights have transferred are "eligible students." Data Use Standards: Key Terms

[35] HIPPA (1996). The HIPAA Privacy Rule standards implemented by the U.S. Department of Health and Human Services address the use and disclosure of individuals' health information—called "protected health information" by organizations subject to the Privacy Rule ("covered entities")—as well as standards for individuals' privacy rights to understand and control how their health information is used. Data Use Standards: Key Terms

[36] CIPSEA, Part of Title V of the E-Government Act of 2002 (Pub. Law 107–347, 116 Stat. 2899, 44 U.S.C. § 101), "permits wage records created for the UI system to be used for statistical and evaluation purposes. H.R.4174 — Foundations for Evidence-Based Policymaking Act of 2017, passed by the House of Representatives in November 2017, would provide clearer direction to allow for these data to be used for program evaluation purposes." (CREC, "Legal Guide to Administrative Data Sharing for Economic and Workforce Development," State Data Sharing Initiative, March 2018, p. 3). Other data that would be matched with wage data might be subject to additional laws and restrictions, such as education data under the Family and Educational Rights and Privacy Act (FERPA). Such laws could restrict the potential benefit of an improved federal wage system, and those laws may need to be changed or procedures adapted to make the most of various, relevant sources for postsecondary and workforce outcomes information.

Data Use and Interpretation

Stakeholders affirmed their priorities to develop safeguards on the interpretation of data, including protecting the dissemination of erroneous findings, creating common data definitions and dictionaries, and sharing the context and assumptions specific to each agency.

There was a consensus among interviewees that other agencies or external entities can misinterpret their data and publicize erroneous conclusions. Stakeholders expressed that data are very complex and tied to internal business practices. Even if well documented by an agency, translating nuances of data as part of broader data interpretation can be a challenge. The issue of data use and interpretation was a concern for both providers and users of data. Stakeholders noted that sharing report/publication drafts before release would help ease this concern.

Agencies are protective about the data entrusted to them and eager to ensure it is used to draw accurate conclusions. As a result, the majority of stakeholders believe that a modernized system must empower stakeholders with approval over data access and use requests.

Stakeholders also reported a need for shared data dictionaries, reference data, and metadata for NCLDS data elements. This is particularly important when data does not cover the full universe of potential cases, as in CFS data, which only includes those who pay for unemployment insurance in NC, leaving out federal employees or self-employed individuals. For example, the importance of contextual understanding is heightened as agencies navigate the unemployment ramifications of COVID19. With new changes to unemployment insurance claims, the context of those guidelines are important in terms of counting claimants impacted via COVID, and the payment amounts. Understanding the decision rules and assumptions applied within each agency was highlighted as well. For example, NCDPI student counts are taken only on census day 20, thus not capturing student population changes that may occur due to dropouts or transfers at other points in the year.

Stakeholders report that gaining contextual understanding will require cultivating cross-agency relationships and expertise. Several interviewees proposed that this could occur as a formalized job responsibility for existing and new staff or through the creation of new positions with dual employment between NCLDS and an individual agency or that rotate through different agencies.

Each agency and NCLDS will need to have technical staff to oversee data management as well as expertise with an understanding of cross-agency data and context. NCLDS governance will need to establish processes for fielding data requests and decision rules for what data can be accessed by whom.

NCLDS Leadership[37]

NCLDS sets ambitious goals for linking data across agencies and over time to enable data-driven decision-making. To reach those goals, NCLDS will need balanced governance structures and effective leadership.

NC stakeholders emphasized that an effective leader of NCLDS will need strong, relationship-building skills to earn the confidence of elected and appointed officials, to build and maintain buy-in, and to convene stakeholders to drive the initiative forward. They reported that an effective leader would operate from a neutral position, ideally outside of participating agencies, to build consensus and

[37] Art of the Possible: Cross-Agency Data Governance Lessons Learned from Kentucky, Maryland, and Washington

momentum. In partnership with a governing board, an effective NCLDS leader would lead outreach efforts to demonstrate the value of NCLDS and to help secure sustainable funding and support from state lawmakers. Interviewees affirmed that an effective NCLDS leader would be needed to balance and meet the needs of the system, contributing agencies and other stakeholders.

Capacity and Sustainability[38]

North Carolina stakeholders agreed that capacity for data management, reporting, and analytics—at partner agencies as well as the eventual hosting entity of NCLDS—would ultimately determine the usefulness and impact of the system. They pointed to the absence of analytical capacity in most state entities, which currently lack the expertise, time, or both, to pursue the lines of inquiry NCLDS would enable.[39]

Stakeholders agreed that, in order to realize a NCLDS vision—to enhance the availability and access of cross-agency actionable intelligence, data, and metrics to support data and knowledge-driven solutions that address complex social problems along the early childhood, education, and workforce continuum—North Carolina will need to dedicate additional resources.

Stakeholders did not view reliance on competitive federal grant funding as a sustainable solution.[40] They felt that reliance on grant funding was likely to result in systems that aligned with grant requirements at the expense of the state's priorities. Instead, they believed recurring state funding for core functions and capacity is vital to the success of NCLDS, although external funding could help pilot new functionality. Interviewees affirmed that NCLDS could be more successful in securing state funds if agencies advocated collectively on behalf of a system housed at a neutral, trusted entity.

Officials from other states described keys to sustainability, and most began with a description of their funding models. For example, in Connecticut, P20WIN was initially funded by grants, but is currently supported only by in-kind resources from partner agencies. Consequently, P20WIN's sustainability is contingent on agencies' continued participation and commitment of resources.

In addition to securing a mix of federal and state financial support for its Kentucky Longitudinal Data System (KLDS), Kentucky has developed a marketing strategy to maintain stakeholder support and ensure widespread use.[41] The KY Center for Statistics (KYSTATS) representatives meet with all Kentucky legislators to tell them about the system's capabilities. They present at state and national conferences and host their own annual conference where data users can share ways that they have leveraged the longitudinal data. This unique model ensures KYSTATS' continued relevance by keeping stakeholders up to date on its convenience and usefulness. Likewise, Georgia offers in-person and web-based training to all of its users to maintain demand and commitment to their LDS, GA-AWARDS.

[38] SLDS Sustainability Planning Guide

[39] Stakeholders considered the LEAD at the NC Department of Commerce and the University of North Carolina System Office to be notable exceptions to this rule.

[40] Sustainability is the capacity to support a system or program over time with sufficient financial and human resources to meet current and future needs. SLDS Glossary

[41] SLDS Sustainability Toolkit and interview with the Kentucky Center for Statistics' business analytics and project management director.

Sustainability is clearly multifaceted, for it is rooted in funding, relationships, buy-in, outreach, legislative support, and governance. North Carolina will have to consider these factors and more as it determines how to best position the building blocks to support a modernized NCLDS.

Section 4. NCLDS Design and Implementation Considerations

NC stakeholders envision an NCLDS that enables evidence-based policymaking, continuous improvement, and performance management across the early childhood, education, and workforce continuum. Achieving that vision, and ensuring that it is sustainable, will require careful deliberation and an intentional, coordinated design of NCLDS data governance, system architecture, and analytical and research capacity. This section presents context and considerations to help NC stakeholders address open questions about NCLDS:

- **Data governance:** How will data contributors make decisions about their shared information assets?
- **System architecture:** What system structure will support NCLDS information and reporting needs?
- **Analytic and applied research capacity:** How can NCLDS develop internal analytic capacity and external research capacity to improve evidence-based policy, continuous improvement, and performance management?
- **Sustainability and impact:** What short- and long-term funding strategies will NCLDS require to deliver value to stakeholders?

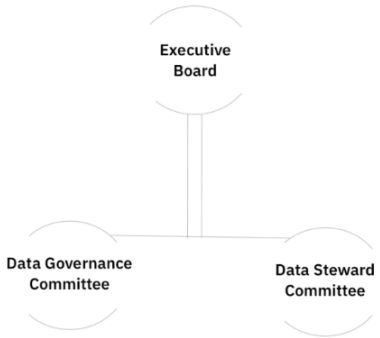
Considerations for Establishing Data Governance and Organizational Structures

Decisions concerning data governance and organizational structures are the key to effective LDS design and implementation. Data governance defines policies and procedures to manage the availability, usability, integrity, quality, and security of data. Data governance establishes clear roles and responsibilities for data contributors and organizes agencies to improve data quality.

Designing the NCLDS Governance Structure[42]

Typically, executive leadership is exercised through an Executive Board (EB), which is composed of representatives from each data-sharing partner, including each entity's senior leadership, chief information or technology officer, or chief data officer.[43] The chairmanship can be permanent or rotating. In Kentucky, the Secretary of Education and Workforce Development is the permanent chair. In Maryland, the chair rotates every four years.

Figure 5: Governance Structure



[42] The SLDS Data Governance Structure Toolkit and other publications offer detailed guidance concerning data governance structures. Comparison states consulted for this project generally conformed with published best practices, with notable nuances concerning organizational affiliations and the number and voting rights of committee members. Due to substantial variability in the number, longevity, and expertise of officials who participated in comparison state interviews and consequent limitations around their ability to assess the effectiveness of their governance structures, this subsection draws more heavily from the SLDS Data Governance Structure Toolkit than from comparison state interviews.

[43] Board membership and chairmanship can be established administratively (in the near term) or legislatively (over the long term).

Executive Board responsibilities include:

- Setting system goals and priorities;
- Approving data governance policies;
- Approving new data partners;
- Ensuring coordination between agencies' data governance programs and the LDS data governance program;
- Modeling and advocating data governance principles with internal and external groups; and
- Holding all staff responsible for adhering to the data policies and processes established through data governance.

In addition to the EB, data governance programs usually have one or two subcommittees, depending on the system's size, staff capacity, and culture. These subcommittees support the EB by providing the program and technical expertise to help guide decisions about the system's structure, operation, and governance.

In a three-group structure, the two subcommittees include a Data Governance Committee (DGC) and a Data Steward Committee (DSC).

The DGC develops and maintains policies and processes for the management and use of cross-sector data. DGC members are empowered by their agency's executive leadership to represent their organization in interagency decision-making. DGC responsibilities include:

- Developing data governance policies;
- Understanding the universe of data contributed by their agencies;
- Modeling and advocating data governance principles with internal staff and other external stakeholders;
- Communicating with the DGC any data issues, regulations, plans, and policies from their program areas that could affect other programs or have an impact on IT; and
- Communicating with their agency staff about the activities and decisions of the DGC.

The DSC comprises staff members and IT representatives from each data contributor with detailed knowledge about their agencies' source systems. DSC responsibilities include:

- Serving as contributing agencies' points of contact for collaboration and coordination of data initiatives, tools, and resources;
- Identifying and escalating issues as necessary to the DGC or EB; and
- Helping to ensure that data quality and timeliness through collaboration are present both within the DSC and with the DGC and others in the agencies.

In short, the DGC sets the context for the data governance program and the DSC helps implement its technical aspects. For a smaller LDS, the DGC and DSC might be combined into a single implementation-level committee.

Once in place, the NCLDS data governance program can adopt policies that set priorities for the system, define acceptable data uses, design processes for requesting and approving data requests, and take steps to integrate governance processes into the daily routines of NCLDS and contributing systems.

During the initial design and implementation phase, the NCLDS data governance program will consider and adopt policies and procedures concerning data scope, data quality, and data access.

Defining the Scope of NCLDS Data

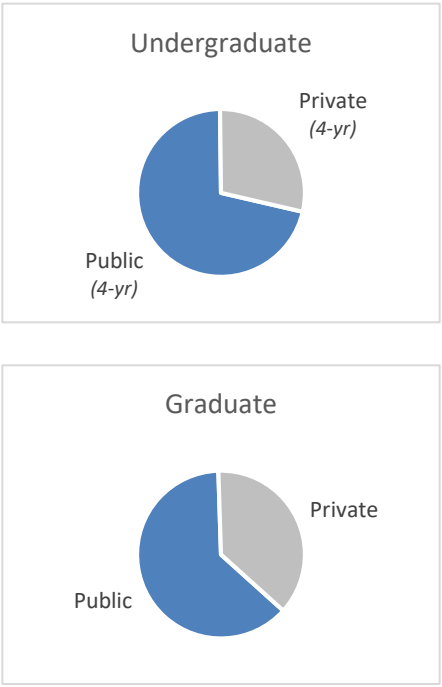
The goals and priorities of NCLDS will shape decisions around agency data sharing and the scope of data available through the system. Interviews with other states suggest that mission- and vision-driven variations in agency participation are common, but most states have core contributors: early childhood, K-12, higher education, and workforce.

NC stakeholders expressed interest in expanding the scope of NCLDS data beyond what ECIDS, NCSW, and CFS provide. For example, several interviewees expressed a preference that NCLDS include data from private colleges and universities in North Carolina.[44]

Among undergraduates enrolled in four-year programs in North Carolina, approximately 190,000 (71%) attend public colleges and universities, while about 77,000 (29%) attend private colleges and universities.[45]

Among graduate students, approximately 47,000 (63%) attend public universities, while about 28,000 (37%) attend private institutions. See Figure 6.[46]

Figure 6: NC Enrollment (Fall 2018)



In terms of additional data sources, several interviewees reported that National Student Clearinghouse data could contribute important information about individuals who are educated out-of-state. Others reported on the value of expanded access to wage data to include federal employees, the self-employed, and NC residents who work out-of-state. These additions could give NCLDS users a more complete picture of how individuals progress along the education to workforce continuum. Other data sources mentioned by NC stakeholders are listed in Appendix 4.

Kentucky and Minnesota offer elements beyond what is commonly available in an educational longitudinal system. Kentucky's KYSTATS includes data from the Cabinet for Health and Family Services (SNAP, TANF, Medicaid eligibility and claims) as well as Justice and Public Safety Cabinet corrections data. Minnesota's system contains data from birth records, economic assistance programs such as SNAP and MFIP, family home visiting data, the Child Care Assistance Program, and the Early Hearing Detection Intervention. Moreover, in Washington State, data from additional sources, such as corrections and

[44] Currently, NCSW does not include data from NCICU member institutions, but CFS is working on a pilot with three institutions to analyze graduates' employment patterns. NCICU has been engaged in the Steering Committee and the NCLDS modernization study and has committed to signing the MOU pending adoption of administrative rules. Among the comparison states, Connecticut and Georgia reported that independent colleges and universities participated voluntarily. Data sharing is mandatory in Kentucky, Maryland, Minnesota, and Washington.
[45] For 2-year programs, public institutions enroll approximately 219,000 individuals, while private institutions enroll about 4,800.
[46] IPEDS North Carolina postsecondary enrollment in four-year colleges and universities.

justice programs, can be linked to education data for specific reports or analyses, but they are not regularly collected in their data warehouse.

NC stakeholders also discussed the importance of deciding how much historical data should be included in an NCLDS. The historical capacity of NCLDS data directly impacts the system's ability to support longitudinal analyses and to answer questions that require analysis over time. NC stakeholders emphasized that NCLDS should be able to answer questions about key transition points, many of which would require data to be retained over a significant period of time. Interviewees offered examples of key questions:

- *What are the impacts of early learning experiences as students move through school and beyond?*
- *What K-12 experiences determine whether a student attends an independent college, community college, or a UNC System school?*
- *What is the impact of extending foster care and adoption benefits to Age 21?*
- *How do workforce outcomes compare for cohorts of certificate, 2-year, and 4-year degree holders?*

Setting NCLDS Data Quality Standards

NCLDS governance will define data quality standards for the system. Data quality standards typically start with consistent, cross-agency data definitions and are reinforced by checks and controls that are agreed upon by contributing agencies and NCLDS governance. Source system adherence to standards will be essential for NCLDS data quality.

The use of consistent data definitions across contributing agencies reduces the burden on agencies to explain their data to other agencies and system users, and it reduces the risk of data misuse or misinterpretation. For data requesters, a publicly available data dictionary—with variable definitions and types, noting the years available, the source system, and other details—simplifies the process of identifying and retrieving the data needed to answer questions about specific policies or programs.

There is a national effort, led by the U.S. Department of Education's Privacy Technical Assistance Center, to develop Common Education Data Standards (CEDS) that align data definitions to facilitate data sharing across states. For example, in North Carolina, NCDPI adopted CEDS as the foundation for its CEDARS warehouse and NCSW implementation. However, the use of CEDS, especially beyond education source systems, is extremely limited. Furthermore, although CEDS can provide some of the functionality of a data asset catalog, North Carolina would need to explore other options to accommodate the broader metadata standards that NCLDS will require.[47]

The governance committee also will need to design checks and controls to ensure that data are accurate, complete, timely, valid, and consistent. For example, CFS has a data cleansing process and business rules for each data element, as well as a workflow process that requires contributing agency

[47] Common Education Data Standards

sign-off to ensure proper use. This process helped identify issues that improved data quality in contributing systems as well.

In an interagency LDS, data quality controls often have to balance competing demands, for example, between completeness and accuracy. Completeness indicates that the data are all there, and accuracy ensures that the data are correct. Accuracy checks may exclude some source system records and result in data being less complete. Setting a high bar for accuracy could reduce the number of records that a longitudinal data system would be able to match across data sources. Understanding the priorities of different user groups will help the NCLDS governance address these and other considerations in determining data quality standards.

Determining NCLDS Data Access Policies (Internal and External)

The NCDIT provides policy requirements around data classification and handling that are outlined in the newly developed NCLDS MOU. Effective data security protocols help safeguard data repositories and protect privacy. NC stakeholders reported that NCLDS could provide in-house data security training for all users, beginning with authorized personnel from data contributors, to ensure a shared understanding of security protocols and expectations for safeguarding data.

NCLDS governance will need to define rules and a process to govern data access.[48] Depending on the process adopted and the characteristics of the requester and the request, the rules could permit automatic approval or require that the appropriate data governance committee vote (unanimously or by majority) to approve or deny the request.

The rules and the process might vary based on a number of factors. For example, characteristics of the requester (if the individual is from a contributing data partner or is internal or external to state government) may indicate a certain level of access. In addition, characteristics of the request itself, for example, which data elements were requested, and the regulations governing the data source, could also trigger specific authorization requirements. In all of these scenarios, rules for small-cell suppression and safeguards around deductive disclosure in public reporting can be used to protect sensitive data with the goal of not compromising the potential of NCLDS to answer important policy questions.

Other states described a range of approaches for submitting and reviewing data requests. Most states document the process online and post application materials that can be submitted via email or an online portal. Both the application requirements and the review process may vary based on the level of data or access requested. In addition, the volume of requests varies substantially between states, for example, in 2019, Kentucky had at total of 239 requests, Washington State had 58 requests, and Connecticut had 21 requests. A governing committee or subcommittee typically reviews these requests and approves or denies them on a case-by-case basis.

For example, in Kentucky, staff follows up with data requestors to discuss the proposed research question as well as the relevant data elements available in KYSTATS. If the data request meets the necessary checks, it is forwarded to a KYSTATS team for approval. In Wisconsin, DPI analysts review the request to be sure that it adheres to agency requirements and is beneficial to the state before deciding whether to recommend it to the Wisconsin Information System for Education (WISE) Steering Committee for approval. In Connecticut, data requests require unanimous approval by its six-member

[48] §116E-4 (a)(4)(c). For a list of current data request processes in NC's existing systems, see Appendix 3.

Data Governing Board and MOAs specify expectations for securing data during transmission and analysis.

In keeping with best practices, states often include multiple checks to ensure that data requests align with the proposed research questions, comply with system guidelines, and execute data sharing agreements following official approval.[49]

NCLDS governance can also design processes to review findings prior to dissemination. In addition to IRB requirements placed on university-based researchers, officials from other states reported inserting specific provisions into data sharing agreements that compel data requesters to submit reports with final results for review prior to publication. Submission periods reported by comparison states ranged from 10 days (Kentucky) to 45 days (Maryland)—with review and notification processes in accordance with state MOUs.

In sum, North Carolina will need to make a series of decisions around governance structures, policies, and processes to build an NCLDS that unleashes the potential of cross-agency, longitudinal data without compromising privacy, security, or trust.

Finding a Home for NCLDS

In considering where an NCLDS should be housed, stakeholders affirmed the importance of neutrality, expertise, reputation among partner agencies, and the ability to secure the confidence of the General Assembly. Although a minority of stakeholders thought NCLDS should be housed in a participating agency such as NCDPI or Commerce, a majority of the stakeholders thought that it should be housed in a neutral entity.[50] Although open to other possibilities, interviewees generally affirmed that GDAC would be the logical choice.[51]

Considerations for Designing the NCLDS Architecture

NCLDS will need to consider the type of system architecture that will best support the goals and priorities of NCLDS users. LDS can be built in different ways—commonly as federated, warehouse, or hybrid systems.

In a warehouse system, a copy of the data from all agencies is integrated and housed in a centralized repository with access granted from a single governing entity. In contrast, a federated system leaves data within the originating agency, where decisions around sharing are made at the agency level for each distinct data request.

A minority of interviewees held strong preferences about design choices. Some focused on system functionality, with the perception that a warehouse supports more efficient reporting and querying, as

[49] Managing Data Requests

[50] The Office of the Governor was also suggested as a potential home for NCLDS by a member of the Steering Committee.

[51] Regardless of which entity ultimately houses NCLDS, stakeholders emphasized that additional dedicated staff would be required to fulfill those responsibilities.

well as the flexibility to more easily scale up to incorporate additional data sources. Others focused on trust and control issues, suggesting that a federated model gives agencies more control over their data, and, therefore, it sidesteps some trust-related roadblocks to collaboration. Similarly, even though interviewees were open to a warehouse approach, some expressed concerns around security and privacy. The primary issue cited around the warehouse data model is concern from individual agencies that they would lose control over their data.

Other state experiences affirm that either design may be used to establish effective and successful LDS.[52] For example, Minnesota and Kentucky built safeguards via tiered access for different user groups to ensure security, regardless of the underlying system model. The existing literature and other state experiences suggest that the two models are not mutually exclusive, and there is an opportunity to build a hybrid design.

The unifying theme across interviewees is that data governance will need to establish clear roles, responsibilities, data stewardship, and ownership regardless of system structure. NCLDS architecture and design will ultimately be dependent on the system characteristics that are recognized as providing the optimal technical structure for partner agencies. Whether data are centrally located or maintained at the source, agencies need to build trust among NCLDS partners and trust in NCLDS data.

Appendix 5 presents attributes of warehouse and federated data models, highlighting perspectives about trade-offs as expressed in stakeholder interviews.

Considerations for Building Analytic (Internal) and Research (External) Capacity

NC will need to determine the balance between internal analytic capacity[53] and external research capacity.[54] Although internal analytic resources are necessary for rapid responses, descriptive analyses, and technical calculations, external research partners may be needed for causal analyses, long-term studies, and the added objectivity provided by independent, external experts (see Figure 7).[55]

The importance of this decision emerged during stakeholder interviews, with an emphasis on the need to add internal analytic capacity. Almost all of the stakeholders reported a gap between existing agency's internal analytic capacity and their interest in answering broader policy research and evaluation questions. Even for agencies that have analytic capacity, current day-to-day programmatic and operational demands can make it difficult for broader policy research, evaluation, and planning to get the attention it deserves within North Carolina.

All of the stakeholders emphasized the need for agency and organization-level expertise in understanding the data. With any analytic system, there will be a need for dedicated staff to help build intra-agency capacity, such as data dictionaries, where they do not yet exist. The key factor is to ensure that the analysts who are doing analytics collaborate with all of the agencies involved.

[52] Early Childhood Integrated Data System Guide

[53] The Case for Government Investment in Analytics

[54] What Research Do State Education Agencies Really Need? The Promise and Limitations of State Longitudinal Data Systems

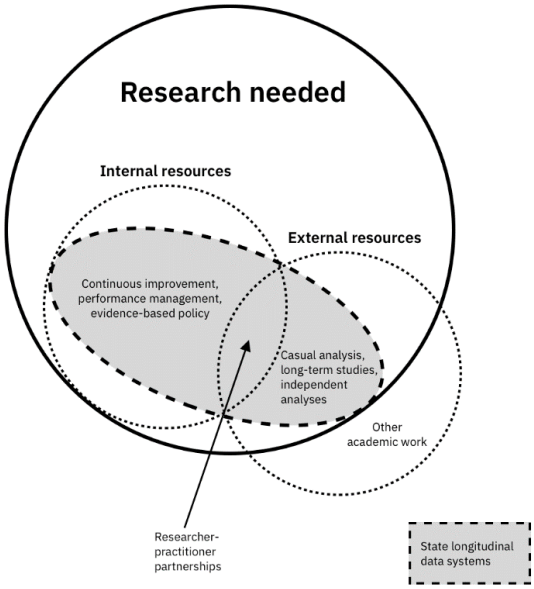
[55] Graphic adapted from What Research Do State Education Agencies Really Need? The Promise and Limitations of State Longitudinal Data Systems

Several interviewees proposed a staffing approach that would involve co-locating and/or rotating data analytic staff within state agencies as a means of building trust and improving collaboration across agencies.

Best practices and lessons from other states stress the importance of building analytic capacity within state government to facilitate policy-related research and evaluation. The data experts who will be building an NCLDS (building data structures and maintaining documentation) will need colleagues with the analytic skills and methodological training necessary to use the longitudinal data for answering relevant research and evaluation questions.

In addition to this analytic capacity, NC can leverage external researchers to use NCLDS and supplemental data sources to perform independent evaluations, causal analyses, and longer-term studies. The literature and state experiences specifically highlight the value of researcher-practitioner partnerships, which are defined as long-term collaborations (between an agency and an external partner), that are organized to investigate problems of practice and generate solutions for improving outcomes.[56]

Figure 7: Research needs supported by LDS



Creating long-term partnerships, rather than short-term and transactional relationships, allows external researchers to increase their understanding of state-level needs; creates space for trust to develop; and enables state staff to explain in what format research may be the most helpful to policymaking efforts.[57]

External researchers are not the only partnerships that NC might consider. Philanthropy can be an important neutral convener, and it has played a role in funding existing, state, evidence-based, policymaking efforts, particularly in the development of a state LDS.

Considerations for Sustainability

Beyond building consensus on governance, design, and capacity, NCLDS success depends on long-term funding and current and future state leaders' commitment to sustaining the program.

Funding

NCLDS will require new funds for developing the technical infrastructure and adequately staffing key system functions. LDS funding options include grant funding and recurring state appropriations, with cost-recovery programs providing a minor, supplemental revenue source in some states. LDS teams that rely heavily on grant funding acknowledge concerns about sustainability. Many states, like Kentucky and Washington State, continue to support their systems with a combination of federal and state funds, while advancing legislation to secure additional state appropriations. Georgia and Maryland's systems

[56] National Network of Research Practice Partnerships

[57] Researcher-policymaker partnerships: Strategies for launching and sustaining successful collaborations.

are sustained largely through state funding. Connecticut, Kentucky, and Maryland report that a small portion of their budgets is funded from fees that are imposed for time-intensive data requests or for requests that are not aligned with their states' research agendas. The NCLDS team could opt to pursue a number of funding strategies, including grants from federal and philanthropic sources, state appropriations, and/or cost recovery programs.

North Carolina interviewees agreed that NCLDS should be funded through recurring state appropriations, for most of the interviewees considered recurring state funds to be an important signal of state prioritization of the system. Although some differed in how funds should be appropriated—whether through agencies or directly to the NCLDS —most of the interviewees considered funds going directly to a neutral entity to be preferred. Interviewees agreed that a joint request for state funding by partner agencies would be the most compelling approach.

Implementation

The costs of what is needed to support and maintain the NCLDS will depend upon decisions about the data system's vision and scope. Interviews with other states did not yield detailed information about LDS design, implementation, maintenance, and operation costs. However, a study by the Data Quality Campaign (DQC) catalogued cost drivers and figures for a handful of states.

According to the DQC report, a data system's costs can be divided into its start-up or implementation costs and its ongoing maintenance costs. Start-up costs tend to vary by system architecture (centralized versus federated) and functionality, by existing linkages and infrastructure, by the number of participating agencies, and by the ease of negotiating data agreements.[58] Maintenance costs are determined in part by the level of demand for data products and analysis, the volume of data requests, hosting expenses, and research capacity.[59]

Figure 8 summarizes the implementation and maintenance costs associated with five state data systems.[60] Decisions about NCLDS functionality, scope, and infrastructure will ultimately determine the work effort and resources required, but this table provides some indications of the scale of LDS implementation and maintenance costs.

Figure 8: LDS Implementation and Maintenance Costs

State	Architecture	Costs
Illinois Illinois Longitudinal Data System (ILDS)	Federated	*Implementation*: (not known) *Maintenance:* $310,000 a year

[58] Costs of State Longitudinal Data Systems. Other start-up cost considerations include vendor vs. in-house development, available funding, when the system was built (many technology costs decrease over time), level of organization and planning, complexity of data governance and policy procedures, and data quality.
[59] Costs of State Longitudinal Data Systems. Other maintenance cost considerations include the technological sophistication of the hardware and software used, and the amount of data and new data acquisitions.
[60] Costs of State Longitudinal Data Systems

Maryland Maryland Longitudinal Data System (MLDS)	Centralized	*Implementation:* $2,747,000 *Maintenance:* $2,077,000/year
Nevada Nevada P-20 to Workforce Research Data System (NPWR)	Federated	*Implementation:* $2,500,000 *Maintenance:* $450,000/year
Utah Utah Data Alliance (UDA)	Centralized	*Implementation:* $7,144,934 *Maintenance:* $1,800,000/year
Virginia Virginia Longitudinal Data System (VLDS)	Federated	*Implementation:* $7,500,000 *Maintenance:* $475,000/year

Demonstrating Impact

NCLDS will need to consider approaches to engaging and communicating with the NCGA, the Office of the Governor, and the public about the benefits of the NCLDS. Articulating the value provided by the system is a powerful way to demonstrate that the NCLDS merits continued support from its stakeholders. The value may include improved data to support policymaking, analyses to support program improvement efforts, cost, and/or resource savings.

Connecticut noted that agencies' interest in their data system's potential to reduce the burden on agency staff for reporting and analysis. Minnesota and Kentucky emphasized that they demonstrated value by answering legislators' data queries. Washington State, Kentucky, Maryland, and Wisconsin reported that consistent stakeholder engagement helped them develop and market services responsive to stakeholders' needs, which enabled them to maintain support among partner agencies and policymakers through leadership changes.[61]

[61] Four best practices for implementing state longitudinal data systems

Section 5. Recommendations

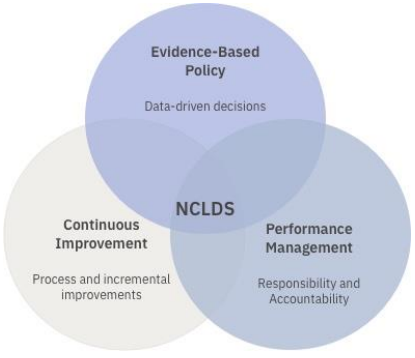
NCLDS should be designed to support evidence-based policymaking, continuous improvement, and performance management. These recommendations reflect our assessment of steps that North Carolina must take to establish a system with the capacity to inform decisions regarding policy and programs along the early childhood, K-12, higher education, and workforce continuum. The recommendations are organized topically (rather than sequentially) into eight categories. Collectively, they put the people and processes in place needed to support NCLDS. These conclusions are based on our interviews with NC stakeholders, discussions with system experts in other states, and a review of best practices from recognized experts such as the U.S. Department of Education's SLDS project, the Data Quality Campaign, and Actionable Intelligence for Social Policy.

NC stakeholders shared a range of perspectives about how trust factors into NCLDS modernization efforts. Trust bolsters relationships between individuals—from system leaders to program and data experts—to improve collaboration across agencies. Our recommendations are designed to nurture trust by structuring stakeholder engagement, governance, analytical and research capacity, and sustainable investments to provide transparency and to facilitate collaboration across partner agencies and entities. The recommendations marked with  are part of a comprehensive strategy to build trust.

Recommendation 1: Affirm the NCLDS vision as a "system of systems" that links data from across agencies and over time to support evidence-based policy, performance management, and continuous improvement:

- Eliminate use of the term ELDS in outreach and education efforts to ensure clarity regarding the relationship between NCSW and NCLDS.
- Use the term NCLDS to identify a "system of systems" that links data across early childhood, K-12, postsecondary education, and workforce. Beginning immediately, employ the term NCLDS consistently in documentation, communication, and related online resources.

Figure 9: NCLDS Vision



- Revisit General Statute 116E to address ambiguity regarding NCLDS definition, location, and governance. NC stakeholders identified the following examples of ambiguity:
  - §116E is titled "Education Longitudinal Data System," although §116E-5 creates the "North Carolina Longitudinal Data System."

- §116E-4 specifies the powers and duties of GDAC, but §116E-5(a) states that the system will be located administratively within the Department of Public Instruction.

- §116E-4(a) directs GDAC to establish a committee on data quality but does not confer authority to establish a governance structure to coordinate participation of partner agencies.

Recommendation 2: Craft a stakeholder engagement strategy to build support for NCLDS:

Leverage Education Cabinet support for NCLDS to secure buy-in from the General Assembly and key system and agency decision-makers.

Engage a broad group of NCLDS stakeholders using an equity approach—policymakers, agency leaders, program managers, data contributors, legislative and fiscal analysts, postsecondary leaders, school districts, parents/families, and community leaders—to participate in NCLDS design and to identify questions that it should be able to answer.[62]

- Encourage contributing partners to promote NCLDS' potential to support data- and knowledge-driven solutions that address policy challenges along the early childhood, K-12, postsecondary education, and workforce continuum.

- Offer user-friendly information sessions for different user groups on NCLDS functionality, including dashboards, analytical tools, and other system outputs.

Recommendation 3: Establish NCLDS governance and organizational structures:[63]

Establish an Executive Board (by statute) comprised of senior leaders or their designees from each contributing agency. In the interim, appoint an Executive Committee (by administrative action) with the same representation.

Hire an NCLDS Executive Director to oversee the implementation and operation of the system and to act as a liaison between the Executive Board/Committee and NCLDS data governance committees.

Name representatives from data-contributing agencies and entities to Data Governance and Data Steward committees.

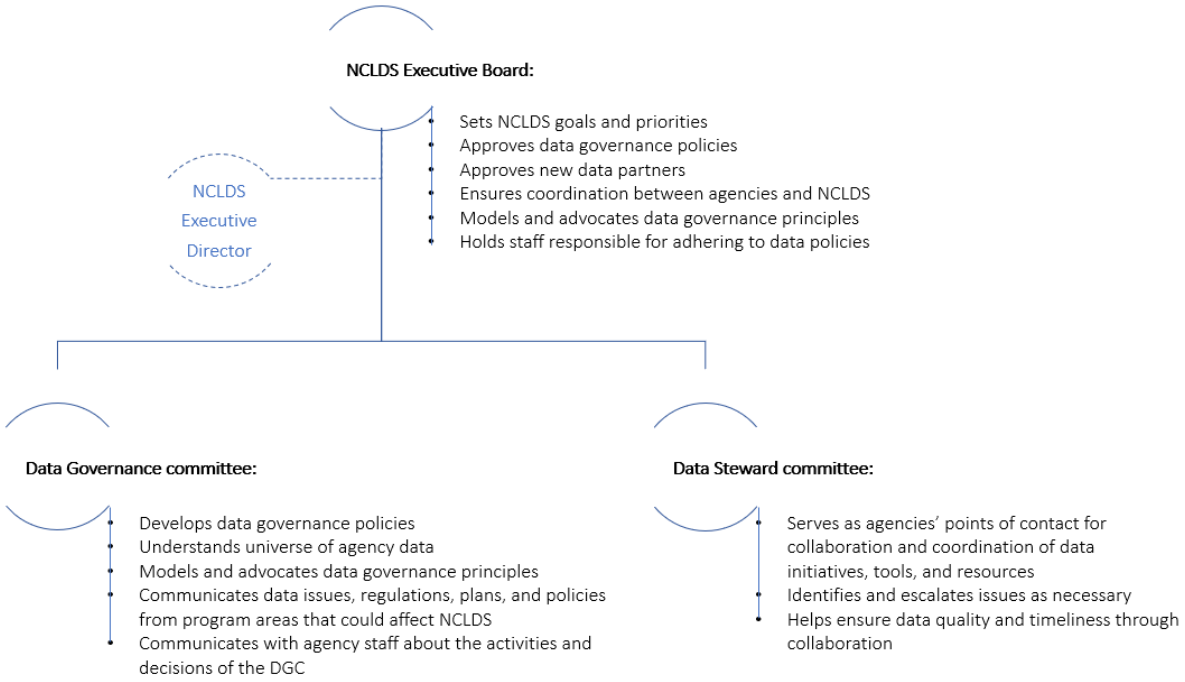
[62] The Hunt Institute's Informed Decision-Making Collaborative, funded by the Bill and Melinda Gates Foundation, has convened North Carolina stakeholders to identify research questions and begin development of a shared research agenda.

[63] This will require clarifying the relationship between the NCLDS governance structure and existing LDS.

- Adopt data governance policies that provide data stewards with control over the use of their data and that build trust in NCLDS as a partner in safeguarding that data.

- Designate GDAC as the administrative home for NCLDS, leveraging its standing as a neutral entity that has secured stakeholder trust and the trust of the NC General Assembly. This role would be consistent with GDAC's responsibilities to support NC's efforts to improve longitudinal data systems.[64]

Figure 10: Proposed NCLDS governance structure



[64] As identified in statute, GDAC's role in statewide data integration and sharing is to identify data integration and business intelligence opportunities that improve the efficiency and effectiveness of state agencies, departments, and institutions (G.S. 143B-1385). GDAC currently hosts ECIDS, NCSW, and CFS. Chapter 116E grants GDAC enumerated powers and duties with respect to operation and oversight of North Carolina's Longitudinal Data System.

Recommendation 4: Prioritize NCLDS requirements for a phased implementation:

- Prioritize NCLDS users, starting with policymakers, agency and program leaders next, and external researchers next.

- Document and prioritize functional requirements for a flexible system of inquiry, including standard and customizable dashboards, reports, data stories, knowledge visualizations, research-ready data sets, metrics, and analytics. Identify "low-hanging fruit" (e.g., research-ready data sets) for initial implementation.

- Identify the current data contributors to ECIDS, NCSW, and CFS as initial data contributors for NCLDS.

- Identify and prioritize additional data sources for inclusion in NCLDS over time (e.g., National Student Clearinghouse, U.S. Census Bureau, U.S. Bureau of Labor Statistics data, NC Families Accessing Services through Technology (NC FAST), Juvenile Justice, or NC licensing board data, see Appendix 4).

Communicate the plan to NCLDS stakeholders.

Recommendation 5: Develop a system architecture to meet NCLDS functional and technical requirements:

- Design a scalable NCLDS architecture and system infrastructure to meet the functional priorities identified by NCLDS governance.

Include NCLDS stakeholders in the system design process to ensure that the technical infrastructure has the flexibility to support NCLDS' short- and long-term goals and the safeguards to address concerns about data quality, data security, and data privacy.

Form a working group to broaden and to formalize discussions about data quality issues that need to be resolved.

Coordinate and further develop data privacy practices and procedures with data contributors.

- Establish a centralized system for entity resolution. In the interim, NCLDS can use source systems' unique identifiers and manual crosswalks. Over the longer term, leverage GDAC's Enterprise Entity Resolution (EER) to provide a consistent and scalable mechanism for linking data and adding data sources over time.[65]

[65] Research and development on EER methodology are already underway.

Recommendation 6: Build analytic capacity to support NC's evaluation and research priorities:

Hire a Director of Analytics and Research early in the process, who will report to the Executive Director, to oversee the internal analytics team, as well as to serve as the interface with external research partnerships. The Director of Analytics and Research would implement a plan based on the priorities set by the NCLDS Executive Board/Committee.

Establish an internal analytics team with subject matter expertise to support evidence-based policymaking, continuous improvement, and performance management.

Implement a staffing strategy that offers training rotations and/or dual employment in NCLDS and contributing agencies and entities to build cross-agency program and data fluency for new and existing analytic staff.

Recommendation 7: Develop an external research agenda to prioritize data requests and the formation of research-practice partnerships:

Create a collaborative, NC (internal) policy research working group composed of agency representatives, subject matter experts, and practitioners to develop a research agenda[66] and to enhance understanding of cross-agency priority areas.

- Design a framework that prioritizes data requests that are aligned with the research agenda, and that supports the formation of (external) research-practice partnerships.

Establish research approval requirements to ensure that (a) originating agencies approve data use and provide aggregated or de-identified data, (b) researchers maintain active IRB approval status and comply with NCLDS research review requirements, and (c) research findings are made available to the public as deemed appropriate by NCLDS governance committees.

- Create research-ready datasets to guide external researchers and applied policy analysts toward priority topics within the research agenda, providing researchers with high-quality data while making more efficient use of staff time than responding to "one-off" requests.

- Form research-practice partnerships to augment the capacity of the state and its agencies to undertake large-scale research and evaluation initiatives.

[66] This group would continue the work started under the Hunt Institute's Informed Decision-Making Collaborative.

Recommendation 8: Develop an action plan to identify immediate next steps and to secure sustainable funding for NCLDS design, implementation, and operation:

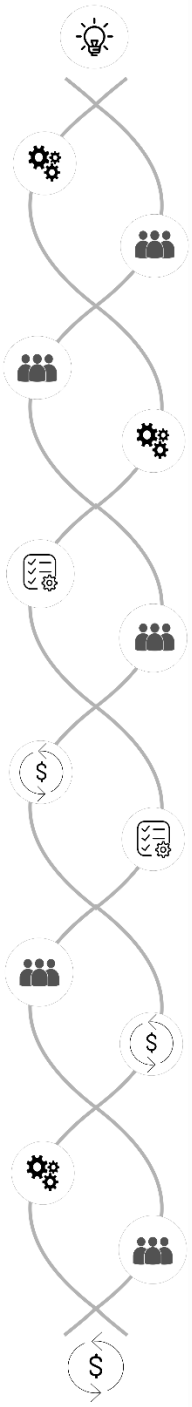
- Move NCLDS forward in the short-term by: (a) asking the Education Cabinet to endorse an action plan, (b) reallocating existing resources to staff key NCLDS positions in the interim, including an Executive Director, (c) establishing the Executive Committee (administratively) and the Data Governance and Data Steward committees, (d) prioritizing functional requirements and identifying "low-hanging fruit" for initial implementation, and (e) estimating the technical and non-technical resources needed from each agency.

- Advocate for state appropriations with support from partner agencies to fund: (a) NCLDS design, (b) initial NCLDS implementation and source system upgrades, (c) system maintenance and operations, and (d) staffing costs for positions dedicated to NCLDS analytic and technical operations.

- Leverage federal and private foundation grant opportunities to support key priorities of stakeholders in piloting new functionality, adding data sources, and forming research-practice partnerships.

A Roadmap to NCLDS

Putting people and processes in place



Out of the gates… **Down the road…**

Affirm the NCLDS vision as a "system of systems"

- Eliminate use of the term ELDS
- Use the term NCLDS to identify a "system of systems"
- Revisit GS116E to address ambiguity

Craft a stakeholder engagement strategy to build support for NCLDS

- Leverage Education Cabinet support
- Engage a broad group of stakeholders to participate in NCLDS design
- Encourage partners to promote NCLDS' potential to support data-driven solutions
- Offer user-friendly information sessions for different user groups

Establish NCLDS governance and organizational structures

- Appoint an Executive Committee (by administrative action)
- Hire an NCLDS Executive Director
- Name representatives from data partners to Data Governance and Data Steward committees
- Designate GDAC as the NCLDS administrative home
- Establish an Executive Board (by statute)
- Adopt data governance policies and procedures that provide data stewards with control over the use of their data and build trust in NCLDS

Prioritize NCLDS requirements for phased implementation

- Prioritize NCLDS users/uses and functional requirements for initial implementation
- Identify current contributors to ECIDS, NCSW, and CFS as initial data contributors for NCLDS
- Communicate plan to stakeholders
- Identify additional NCLDS users/uses, functional requirements, and data sources for expansion over time

Develop a system architecture to meet NCLDS technical and functional requirements

- Design a scalable NCLDS architecture
- Form a data quality working group
- Coordinate data privacy practices and procedures
- Use crosswalks of source systems' unique identifiers for interim entity resolution
- Leverage Enterprise Entity Resolution to provide a consistent and scalable mechanism for linking data and adding new data sources

Build analytic capacity to support NC's evaluation and research priorities

- Hire a Director of Analytics and Research
- Establish an internal analytics team
- Implement a staffing strategy that offers training rotations and/or dual employment to build cross- agency program and data fluency

Develop an external research agenda to prioritize data requests and research-practice partnerships

- Create a collaborative NC (internal) policy research working group to develop a research agenda
- Design a framework that prioritizes data requests aligned with the research agenda
- Establish research approval requirements
- Create research-ready datasets
- Form research-practice partnerships

Develop an action plan to identify next steps and address NCLDS funding and sustainability

- Reallocate existing resources to staff interim NCLDS positions and estimate additional resources needed from each agency
- Advocate for state appropriations with support from partner agencies to fund NCLDS
- Leverage federal and foundation grant opportunities to pilot new functionality, add data sources, and establish research-practice partnerships

References

Brown, J. L. (2019). *Four best practices for implementing state longitudinal data systems*. Education Dive. https://www.educationdive.com/news/4-best-practices-for-implementing-state-longitudinal-data-systems/552355/

Common Education Data Standards. (n.d.). https://ceds.ed.gov/

Conaway, C., Keesler, V., & Schwartz, N. (2015). What research do state education agencies really need? The promise and limitations of state longitudinal data systems. *Educational Evaluation and Policy Analysis*, *37*(1_suppl), 16S-28S. https://journals.sagepub.com/doi/pdf/10.3102/0162373715576073

Connecticut

P20 WIN. https://www.ct.edu/p20win

P20 WIN Data Governance Policy (version 2.0). https://www.ct.edu/files/pdfs/AttachmentD_P20WIN_DataGovernancePolicy_Signed_052815.pdf

P20 WIN Data Governance Manual (version 4.0). https://www.ct.edu/files/pdfs/P20WIN-DataGovManualFinal-061515.pdf

P20 WIN Data Request Management Procedure. https://www.ct.edu/files/pdfs/P20WIN-DataRequestProcedure-Final_01202015.pdf

Data Quality Campaign. (2018). *The art of the possible: cross-agency data governance lessons learned from Kentucky, Maryland, and Washington. https://2pido73em67o3eytaq1cp8au-wpengine.netdna-ssl.com/wp-content/uploads/2018/03/DQC-Cross-Agency-Gov-CaseStudy-032218.pdf*

Data Quality Campaign. (n.d.) *Safeguarding data.* https://dataqualitycampaign.org/topic/safeguarding-data/

Fantuzzo, J., Henderson, C., Coe, K., Culhane, D. (2017). *The integrated data system approach: A vehicle to more effective and efficient data-driven solutions in government.* actionable intelligence for social policy, University of Pennsylvania. https://1slo241vnt3j2dn45s1y90db-wpengine.netdna-ssl.com/wp-content/uploads/2017/09/The-IDS-Approach_Fantuzzo-et-al.-2017_Final.pdf

Georgia

GA•AWARDS. https://gosa.georgia.gov/report-card-dashboards-data/statewide-longitudinal-data-system-gaawards
Data Dashboards. https://gosa.georgia.gov/report-card-dashboards-data/data-dashboards
Data Requests. https://gosa.georgia.gov/report-card-dashboards-data/data-requests
Downloadable Data. https://gosa.georgia.gov/report-card-dashboards-data/downloadable-data

Kentucky

KLDS. https://kystats.ky.gov/
KLDS Data Request Process. https://kystats.ky.gov/Reports/DataRequest

Kentucky Center for Education and Workforce Statistics. (2017). *Data access and use policy.* https://kcews.ky.gov/Content/DataAccessAndUsePolicy.pdf

Labor and Economic Analysis Division. (2018). *A Report on the operations of the North Carolina Common Follow-up System.* North Carolina Department of Commerce. https://files.nc.gov/nccommerce/documents/LEAD/Common-Followup-System/CFS_Operational_Report_2018.pdf

Labor and Economic Analysis Division. (2019). *A Report on the operations of the North Carolina Common Follow-up System.* North Carolina Department of Commerce. https://nccareers.org/CFS/reports/CFS_Operational_Report_May_2019.pdf

Labor and Economic Analysis Division. (2020). *A Report on the operations of the North Carolina Common Follow-up System.* North Carolina Department of Commerce. https://nccareers.org/CFS/reports/CFS_Operational_Report_May_2020.pdf

Leventoff, J. (n.d.). *Costs of longitudinal data systems.* Workforce Data Quality Campaign. https://m.nationalskillscoalition.org/resources/publications/file/Cost-of-State-Longitudinal-Data-Systems_web.pdf

Maryland

MLDS. https://mldscenter.maryland.gov/
MLDS Data Inventory. https://mldscenter.maryland.gov/DataInventory.html
MLDS Research Agenda. https://mldscenter.maryland.gov/ResearchAgenda.html

Minnesota

ECLDS. http://ecIds.mn.gov/
SLEDS. http://sleds.mn.gov/
SLEDS Data Access and Management Policy (2018). https://education.mn.gov/mdeprod/idcplg?IdcService=GET_FILE&dDocName=MDE034706&RevisionSelectionMethod=latestReleased&Rendition=primary
SLEDS Data Access Request Process and Timeline (2019).

https://education.mn.gov/mdeprod/idcplg?IdcService=GET_FILE&dDocName=033954&RevisionSelectionMethod=latestReleased&Rendition=primary
SLEDS Data Interpretation Guide (2018). https://education.mn.gov/mdeprod/idcplg?IdcService=GET_FILE&dDocName=059077&RevisionSelectionMethod=latestReleased&Rendition=primary
SLEDS Scoring Template and Rubric (2017). https://education.mn.gov/mdeprod/idcplg?IdcService=GET_FILE&dDocName=033955&RevisionSelectionMethod=latestReleased&Rendition=primary
SLEDS Governance Charter (2015). https://education.mn.gov/mdeprod/idcplg?IdcService=GET_FILE&dDocName=MDE087818&RevisionSelectionMethod=latestReleased&Rendition=primary

National Center for Education Statistics. (n.d.). *Traveling through time: The forum guide to longitudinal data systems.* Book 3: Effectively Managing LDS Data. https://nces.ed.gov/forum/ldsguide/book3/ch_8.asp

National Network of Research Practice Partnerships. (n.d.). http://nnerpp.rice.edu/

Owen, J. W., & Larson, A. M. (2017). *Researcher-policymaker partnerships: Strategies for launching and sustaining successful collaborations* (1st ed.). New York, NY: Routledge. doi:10.4324/9781315212722

Rhode Island
DataHUB. http://ridatahub.org/
Data Dictionaries. http://ridatahub.org/dictionary/
Data Stories. http://ridatahub.org/datastories/

SLDS Grant Program. (2014). *Data quality: Striking a balance.* SLDS Issue Brief. Institute for Education Sciences. https://slds.grads360.org/services/PDCService.svc/GetPDCDocumentFile?fileId=5044

SLDS Grant Program. (2016). *Identifying SLDS users and their information needs.* SLDS Issue Brief. Institute for Education Sciences. https://slds.grads360.org/services/PDCService.svc/GetPDCDocumentFile?fileId=23252

SLDS Grant Program. (2016). *Sources and linking strategies for employment data.* SLDS Issue Brief. Institute for Education Sciences. https://slds.grads360.org/services/PDCService.svc/GetPDCDocumentFile?fileId=20559

SLDS Grant Program. (2017). *P-20W+ data governance: Tips from the states.* SLDS Best Practices Brief #4. Institute for Education Sciences. https://slds.grads360.org/services/PDCService.svc/GetPDCDocumentFile?fileId=25962

SLDS Grant Program. (2019). *Managing data requests*. SLDS Topical Webinar Summary. https://slds.ed.gov/services/PDCService.svc/GetPDCDocumentFile?fileId=34570

SLDS Grant Program. (2020). *Best practices for the design and implementation of data privacy and security programs.* SLDS Issue Brief. Institute for Education Sciences. https://slds.grads360.org/services/PDCService.svc/GetPDCDocumentFile?fileId=36037

SLDS Grant Program. (2020). *How to engage and train stakeholders regarding privacy and security best practices.* SLDS Issue Brief. Institute for Education Sciences. https://slds.grads360.org/services/PDCService.svc/GetPDCDocumentFile?fileId=35215

SLDS State Support Team. (n.d.) *SLDS early childhood integrated data system guide*. https://childcareta.acf.hhs.gov/sites/default/files/public/slds_ec_integrated_data_system_guide.pdf

SLDS State Support Team. (2019). *SLDS glossary.* https://nces.grads360.org/services/PDCService.svc/GetPDCDocumentFile?fileId=35150

SLDS State Support Team. (n.d.). *SLDS Sustainability planning guide.* https://nces.ed.gov/programs/slds/pdf/sustainability_guide.pdf

SLDS State Support Team. (n.d.). *SLDS sustainability toolkit.* https://nces.ed.gov/programs/slds/pdf/sustainability_toolkit.pdf

SLDS Target Team. (2014). *The framework for data systems*. Institute for Education Sciences. https://nces.ed.gov/programs/SLDS/pdf/Framework_for_Data_Systems_May2014.pdf

SLDS Technical Assistance Program. (n.d.). *Data governance toolkit: structure.* Institute of Education Sciences. https://slds.ed.gov/#program/data-governance-structure

SLDS Technical Assistance Program. (n.d.). *Data use standards: Professional behaviors.* Institute of Education Sciences. https://slds.ed.gov/?scrollTo=b1a#program/data-use-standards:-professional-behaviors

SLDS Technical Assistance Program. (n.d.). *Data use standards: Key terms.* Institute of Education Sciences. https://slds.ed.gov/#program/data-use-standards:-key-terms

U.S. Census Bureau. (2017). *Data stewardship executive policy committee.* https://www2.census.gov/foia/ds_policies/ds017.pdf

Washington State ERDC. https://erdc.wa.gov/

Data Governance. https://erdc.wa.gov/about-us/data-governance
Data Request Process. https://erdc.wa.gov/data-resources/data-request-process
Research Priorities. https://erdc.wa.gov/about-us/research-priorities

Wisconsin

ECIDS. https://dpi.wi.gov/early-childhood/ecids
WISE. https://dpi.wi.gov/wise
WISE Data Elements. https://dpi.wi.gov/wise/data-elements
WISE Data Privacy. https://dpi.wi.gov/wise/data-privacy
WISE Data Requests. https://dpi.wi.gov/wise/data-requests

Wiseman, J. (2019). *The case for government investment in analytics.* Ash Center for Democratic Governance and Innovation, Harvard University Kennedy School. https://datasmart.ash.harvard.edu/news/article/case-government-investment-analytics

Appendices

Appendix 1: North Carolina stakeholder interviews

Executive

Governor's Office
- Geoffrey Coltrane, Senior Education Advisor

North Carolina Division of Information Technology
- Carol Burroughs, Director of Analytics, GDAC
- John Correllus, Deputy State CIO, Chief Data Officer, and Director of GDAC
- Tracy Doaks, Secretary and State Chief Information Officer
- Tom Frantz, former Applications Systems Analyst, GDAC
- Jacqueline Keener, Assistant Director of Analytics, GDAC
- Anna Szamosi, Deputy General Counsel for Data Privacy
- Jessica Wilkins, Business Intelligence Specialist, GDAC

Office of State Budget and Management
- Brett Altman, Budget Analyst
- James Aughenbaugh, Budget Analyst
- Anthony Brito, Budget Analyst
- Erin Matteson, Assistant State Budget Officer
- Jennifer Neisner, Budget Analyst
- Charles Perusse, State Budget Director
- Kevin Rich, Budget Analyst
- James Robinson, Budget Analyst
- Gary Thomas, Assistant State Budget Officer

North Carolina Department of Commerce
- Meihui Bodane, Assistant Secretary for Policy, Research & Strategy, LEAD
- Elizabeth Crabill, Chief Deputy Secretary, Office of the Secretary
- Beth Gargan, Deputy Secretary of Communications and External Affairs, Division of Workforce Solutions
- Sharon Johnston, Attorney, Division of Employment Security
- Elizabeth McGrath, Director of Workforce Research and Evaluation, LEAD
- Oleksandr Movchan, Director of Data Analytics and Research, LEAD

North Carolina Department of Health and Human Services
- Mandy Cohen, Secretary, Department of Health and Human Services
- Rebecca Planchard, Senior Early Childhood Policy Advisor, Secretary's Office
- Hayley Young, Early Childhood Data Manager

Legislative

North Carolina General Assembly Legislative Analysis Division
- Virginia Barlow, Research Assistant
- Brian Gwyn, Staff Attorney and Legislative Analyst
- Kara McCraw, Staff Attorney and Legislative Analyst
- Samantha Yarborough, Staff Attorney and Legislative Analyst

North Carolina General Assembly Fiscal Research Division
- Stephen Bailey, Fiscal Analyst
- Erin Biggers, Fiscal Analyst
- Lisa Fox, Fiscal Analyst
- Deborah Landry, Fiscal Analyst
- Luke MacDonald, Fiscal Analyst
- Eric Moore, Fiscal Analyst

K-12 Education

State Board of Education
- Eric Davis, Chairman, State Board of Education
- Deanna Townsend-Smith, Director of Operations and Policy

North Carolina Department of Public Instruction
- Diane Dulaney, NC SchoolWorks Coordinator
- Chloe Gossage, Chief Strategy Officer
- Karl Pond, Enterprise Data Manager, Business Director, Data, Research, and Federal Reporting

Higher Education

University of North Carolina System
- Andrew Kelly, Senior Vice President for Strategy and Policy
- Gianna Malak, Lead Business Intelligence Analyst
- Diane Marian, Interim Vice President for Data and Analytics
- Kim van Noort, Senior Vice President for Academic Affairs and Chief Academic Officer

North Carolina Independent Colleges and Universities
- Vicki Humphreys, Director of Information Technology
- Tom West, Vice President for Government Relations and General Counsel
- Hope Williams, President

North Carolina Community College System
- Peter Hans, President
- Jennifer Haygood, Chief of Staff/Executive Vice President
- Bill Schneider, Associate Vice President of Research and Performance Management
- Ashley Sieman, Director of Analytics and Reporting

Additional Entities

myFutureNC
- Cecilia Holden, CEO and President

Carolina Demography
- Rebecca Tippet, Director

The Hunt Institute
- Allison Goff, Policy Analyst
- Cheryl Krohn, Senior Policy Analyst
- Patrick Sims, Director of Policy and Research

Appendix 2: Experts consulted from comparison states and national organizations

- **Connecticut**
  Program Manager, P20WIN
  Chief Performance Officer, Connecticut State Department of Education

- **Georgia**
  (*former*) Vice President for Research, Georgia Independent College Association

- **Kentucky**
  Business Analytics and Project Management Director, Kentucky Center for Statistics
  (*former*) Executive Director, Kentucky Center for Education and Workforce Statistics*

- **Maryland**
  Director, Research Maryland Longitudinal Data System*

- **Minnesota**
  Manager, Research and SLEDS, Minnesota Office of Higher Education
  Agency Performance Manager, Minnesota Department of Employment and Economic Development

- **Rhode Island**
  Research Specialist, Rhode Island Department of Education Office of Data and Technology Services

- **Washington**
  Project Manager, Washington P20W Statewide Longitudinal Data System
  Director, Washington P20W Statewide Longitudinal Data System
  Senior Forecast Analyst, Washington P20W Statewide Longitudinal Data System

- **Wisconsin**
  Research Analyst, Policy and Budget Team, Wisconsin Information System for Education
  Associate Director for Programs and Management, Institute for Research on Poverty

- **Data Quality Campaign**
  Director, Policy and Research Strategy
  Director, Policy and Advocacy

- **State Longitudinal Data Systems Research Project**
  Research Director

- **Statewide Longitudinal Data Systems Grant Program**
  State Support Team Program Director and Technical Assistance Lead*

- **U.S. Department of Education, Institute of Education Sciences**
  Program Officer, National Center for Education Statistics

**Presented and responded to questions as part of the Hunt Institute's IDMC session.*

Appendix 3: Current data request approval processes

NCDPI: NCDPI offers a substantial amount of aggregate education data online. Data request forms and other related documents are also posted on the NCDPI website.

Commerce: The Department of Employment Security requires completion of a Data Request form for those seeking access to confidential Unemployment Compensation information.

NCDHHS: NCDHHS is committed to transparency and data sharing for purposes of meaningful research and monitoring to the greatest extent allowed under state and federal laws. Specific rules may apply to specific datasets, but in general, aggregate or de-identified data may be made available for secondary use. Would-be data recipients are asked to fill out a Data Request Form prior to obtaining data. In the future, the Department intends to explore differentiated data request standardized processes for internal requests made by employees of state government.

NCCCS: NCCCS responds to ad hoc requests for aggregate data but does not have a standard practice. NCCCS also works with certain research centers on grant-related projects that leverage record-level data. All use of record-level data research requires an MOU and a Data Sharing Agreement.

NCICU: NCICU, at this time, does not share data for research purposes. NCICU shares statistical data with member institutions only.

UNC System: The UNC System Office does not have official guiding principles regarding data sharing. Data requests are evaluated on a case-by-case basis. Requestors must complete a Data Request form and execute an MOU and a Data Sharing Agreement.

Appendix 4: Additional data sources

The **American Community Survey** is an annual survey of 3.5 million American households that collects data about jobs and occupations, educational attainment, insurance coverage, disability, home ownership, and other topics.

The **Bureau of Labor Statistics** measures labor market activity, working conditions, price changes, and productivity in the U.S. economy to support public and private decision making.

FEDES (Federal Employment Data Exchange System) is an initiative that provides Federal and state agencies with an effective way to include Federal employment information in support of performance and evaluation reports required by Federal and state regulations. Quarterly data exchange is supported between all 50 states and two Federal agencies: the Office of Personnel Management (OPM) and the Department of Defense. There are currently 41 states and the District of Columbia participating in FEDES. The FEDES grant is in the final stages of transitioning to the Kansas Department of Commerce.

Food and Nutrition Services (Food Stamps) is a federal food assistance program that provides low-income families the food they need for a nutritionally adequate diet.

Juvenile Justice (in the North Carolina Department of Public Safety) serves at-risk youth through Community Programs, Court Services, Facility Operations, Education Services and Clinical Services and Programs.

LEHD (Longitudinal Employer-Household Dynamics) maintains an active research program oriented on the use of longitudinally linked employer-employee data.

National Student Clearinghouse contains data on 97%, or 19.5 million, of currently enrolled postsecondary students, from 99% of all public and private postsecondary institutions.

NC FAST (North Carolina Families Accessing Services through Technology) allows Subsidized Child Care Assistance providers to enroll in the program, accept children into their care, update rates for services, and record attendance.

NC Medicaid serves over two million beneficiaries by providing health insurance coverage for North Carolinians with low income, funding necessary services for people with severe mental health needs, supporting in-home and facility-based services in adult care and nursing homes, supporting medically fragile children through its Community Alternatives Program for Children, and assisting children and adults with developmental disabilities.

NC Pre-K is designed to provide high-quality educational experiences to enhance school readiness for eligible four-year-old children.

North Carolina State Center for Health Statistics (SCHS) has state and county level data for a wide variety of health topics including births, mortality, disease prevalence, health behaviors, and more.

Professional certifications from licensing boards, professional associations, and private credentialing entities.

SWIS (State Wage Interchange System) is a mechanism through which States can exchange wage data on an interstate basis with other states needing it for WIOA performance reporting.

Work First, North Carolina's Temporary Assistance for Needy Families (TANF) program, provides parents with short-term training and other services to help them become employed and to move toward self-sufficiency.

Appendix 5: System architecture considerations

	Warehouse	Federated
Data Stewardship	Source agency retains stewardship over its data.	Source agency retains stewardship over its data.
Implementation Time	Requires the design of the centralized warehouse and extraction, transfer and load (ETL) process	Dependent on the implementation timeline at each agency
Data Workload	More staff work up front to perform ETL. With higher utilization, this leads to a lower workload/request	Staff work is largely on a per data request basis. For lower utilization, this leads to a lower workload/request
Data Extraction	Extraction requires match once, use many times	Extraction requires match for each use
Security	Higher stakes in the event of a system breach since all the data are centrally housed	Greater opportunity, but lower stakes, in the event of a system breach. More connections so more possible failure points
Data Quality	Uniform ETL process can provide a consistent level of data quality.	Data quality is dependent on the processes at each agency
Data Retention	Historical data can be maintained centrally, so no additional burden on contributing systems to increase their length of data retention.	Contributing systems have to maintain historical data to meet data retention policies.
Usability	Data are all in one place and is snapshot in time. Facilitates data mining and report generation/dashboarding.	Data pulls are of most current data from contributing systems. If additional years of data are needed for a given cohort, the entire data set will need to be recreated.
Trust	Requires trust in the governance process ensuring agencies retain control over their data.	Sidesteps short-term trust concerns because agencies directly fulfill each data request.

Appendix 6: Summary of comparison state findings[67]

	Connecticut	Georgia
LDS Name	**P20 WIN** Preschool through Twenty and Workforce Information Network	**GA•AWARDS** Georgia's Academic and Workforce Analysis and Research Data System
Established by	Statute/Legislation	Federal SLDS Grant
LDS Statute	Conn. Gen. Stat. § 163-10-10a	GA SLDS Grants
Structure	Federated	Centralized
Governance	P20 WIN Executive Board: 7 members, sector commissioners/presidents to set strategic direction and secure needed resources. Data Governing Board: 6 members, some overlap, but mostly performance/planning officers to define, integrate, store, access, report data and also review data requests. Data Steward Committee: 9 members, generally an IT and a data user/researcher from each sector to lend subject matter and technical expertise to the governance process.	Governed by the Alliance of Education Agency Heads' Data Management Committee.
Organizational Home	Connecticut State Colleges and Universities is the administrative lead agency. No centralized data storage. Agencies maintain their own data.	GA•AWARDS is housed within the Governor's Office of Student Achievement.
Data Providers	Board of Regents for Higher Education CT Department of Education CT Conference of Independent Colleges CT Department of Labor CT Office of Early Childhood CT Office of Policy and Management University of Connecticut	Bright from the Start: Department of Early Care & Learning GA Department of Education GA Department of Labor GA Independent College Association GA Professional Standards Commission GA State Charter Schools Commission GA Student Finance Commission Governor's Office of Student Achievement Technical College System of GA University System of GA

[67] Due to substantial variation in the extent of published documents as well as the number, longevity, and expertise of state officials who participated in interviews, cells were left blank when (1) information could not be found (or conflicting information could be found but not verified) online and/or (2) officials were unable to address certain topics as a result of limited tenure, purview, and/or interview time/scope. The authors accept full responsibility for any errors or omissions.

	Connecticut	Georgia
Independent Colleges and Universities	Voluntary. Interest came from ICUs. CT Conference for Independent Colleges acts as a conduit for contacting colleges for data requests. For matching, ICUs preferred P20WIN's process with the Department of Labor over the Office of Higher Education; this choice made it easier to get them on board.	Voluntary. About 20 of the 25 ICUs in the state participate. ICUs participated in development of RQs and data elements, how system would work, frequency of data feeds, types of canned reports, internal review process, etc. It was a challenge to get all MOUs developed/signed but ICUs took the lead because they were the most concerned about being protected. MOU prevents publication of institution names and rankings.
Vision Statement	Connecticut's Preschool through Twenty and Workforce Information Network (P20 WIN) informs sound educational policies and effective educational program practices through the secure sharing of critical longitudinal data across the Participating Agencies to ensure that individuals successfully navigate educational pathways into the workforce.	GA•AWARDS or Georgia's Academic and Workforce Analysis and Research Data System is the state's Pre-K through workforce (P20W) longitudinal data system. It is comprised of education and workforce data from participating state agencies with the goal of linking and providing meaningful and actionable education and workforce data to participating agencies that support research and informed decision-making.
Example Use Cases	What high school academic experiences are the best predictors of students' success? Which indicators can be used to support students to be college and career ready? Which certificates and degrees enable students to work in Connecticut?	What are the characteristics of high school students who are in the labor force? Does the impact of math and science course-taking patterns on college preparedness differ by race/ethnicity and/or socioeconomic status? How does student mobility, between institutions, type of institution, and/or in and out of the workforce, during postsecondary education shape postsecondary outcomes as well as subsequent employment?
Research Agenda	Yes, established during implementation. Key topics: successful transitions, college readiness, student success, effectiveness of teacher education, affordability and sustainability of higher education, innovation and economic growth, and equity.	Key research topics and advocacy areas include: (1) effectiveness of educator preparation programs; (2) effectiveness of strategies and interventions implemented within the State, and (3) educational background of students who experience the least difficulty in transitioning to college. About 70 pre-approved RQs are posted online. These get priority as well as those aligned with the state's strategic goals.
Leadership		
Primary Stakeholders	Researchers, policymakers; most attention goes to agencies and school districts	Partner agencies

	Connecticut	Georgia
Stakeholder Engagement	Initially held conference calls and webinars to talk about data needs; have also gone to other policy centers, UConn, etc. Largely ad hoc, but there are regular governance meetings, webinars, emails, and newsletters. Most focus is on research/policy applications, but districts are also important stakeholders. P20WIN does regular reports for districts as well as the State Board.	
Size of Staff	No dedicated staff	
Data Matching	Data contributors send files with fake IDs and fields for matching to Department of Labor, which matches and produces a key based on fuzzy and probabilistic matching.	GA•AWARDS uses defined rules and trust values to assign a unique identifier to track students across agencies and across time. All personally identifiable information (PII) is removed immediately after the unique ID is applied.
Data Request Process	Requests go through the Data Governing Board; unanimous vote required for decisions of Executive Board and Data Governing Board; template MOA specifies expectations for securing data during transmission and analysis.	Executive Researcher Committee reviews requests. Online rubrics for feasibility, security, substantive review. Not many denied, more typical to have a revise and resubmit. At this time, only researchers from the partner agencies are allowed access.
Data Request Volume	21 requests in total, the vast majority from agencies	
Research Review Process	Yes, via the Data Governing Board. Each agency whose data is included has an opportunity to review and provide feedback on all reports prior to publication.	Yes, embargo until reviewed
Trust	Equity to all participating agencies in governance; data housed at individual agencies.	
Sustainability	No recurring state funds. Key: Good relationships and equity among participating agencies, as well as provision of useful outputs to encourage continued participation.	GA•AWARDS received full state funding to continue its work at the end of RTT.
Key Takeaways for NC	Had written an initial grant for a centralized system but had to redirect as partners were not ready and that system would not work in the CT context.	Participation of ICUs and the fast-tracked list of RQs.

	Kentucky	Maryland
LDS Name	**KLDS** Kentucky Longitudinal Data System	**MLDS** Maryland Longitudinal Data System
Established by	Statute/Legislation	Statute/Legislation
LDS Statute	Ky. Rev. Stat. Ann. § 151B.132	Md. Code, Com. Law § 190-24-701
Structure	Centralized	Centralized
Governance	Governance Board: Permanently chaired by the Secretary of Education & Workforce Development Cabinet. Membership includes leadership from partner agencies: Education and Workforce Development Cabinet, Department of Education; Council on Postsecondary Education; Higher Education Assistance Authority; Cabinet for Health and Family Services.	Governing Board: 12 members, some mandated, some rotating; chair rotates every 4 years. Data Governance Advisory Board: is responsible for data quality and documentation. Research & Policy Advisory Board: manages the research agenda and makes recommendations re: data requests.
Organizational Home	The KY Center for Statistics (KYSTATS) and its Board are attached to the Education and Workforce Development Cabinet, Office of the Secretary. KYSTATS maintains KLDS. The KLDS Executive Director is nominated by the KYSTATS Governing Board and appointed by the Governor.	Data developed and maintained by MLDS Center, an independent state agency. The MLDS Executive Director reports to the Governing Board and oversees the Research & Policy Advisory Board and the Data Governance Advisory Board.
Data Providers	KY Cabinet for Health and Family Services KY Department of Education KY Council on Postsecondary Education KY Education Professional Standards Board KY Education & Workforce Development KY Higher Education Assistance Authority	MD Department of Education MD Department of Labor, Licensing and Regulation MD Department of Juvenile Services MD Higher Education Commission MD Motor Vehicle Administration US Census
Independent Colleges and Universities	Mandatory. All colleges and universities, including Independent Colleges, submit data to the Council on Postsecondary Education, which in turn submits most of those data to KYSTATS.	Mandatory. ICUs must share data with MHEC which feeds data to MLDS.
Vision Statement	The Kentucky Center for Statistics (KYSTATS) collects and links data to evaluate education and workforce efforts in the Commonwealth. This includes developing reports, responding to research requests, and providing statistical data about these efforts so policymakers, practitioners, and the general public can make better informed decisions.	The mission of the MLDS Center is to develop and maintain a data system that contains student data from all levels of education and workforce data in order to provide analyses, produce relevant information, and inform choices to improve student and workforce outcomes, while ensuring the highest standards of system security and data privacy.

	Kentucky	Maryland
Example Use Cases	How many jobs will Kentucky need to fill during the next five years? What will those jobs pay? Do we have the right mix of skills and workers to fill those jobs? Who will educate Kentucky P-12 students in the coming years? What are they specializing in? Are Kentucky students successfully graduating from high school, earning credentials, and gaining employment?	What is the impact of early childhood education experiences and programs on children's school readiness and K-12 outcomes? What happens to students who start at community colleges and do not go on to 4-year institutions?
Research Agenda	Yes, established biennially; has occasionally led to new data being added; request fulfillment prioritized based upon agenda.	Yes; requests must fit within the agenda. 21 questions in 4 categories: K12 readiness, postsecondary readiness and access, postsecondary completion, workforce outcomes.
Leadership	Governor Steve Bashear (2007-2015) was strong supporter of legislation that led to creation of KYSTATS in 2012.	Governor Martin O'Malley (2007-2015) led efforts to collect and use data to improve student achievement, establish a statewide vision, and work collaboratively to develop a plan to build and use statewide longitudinal data systems. Led to creation of MLDS Center.
Primary Stakeholders	Agencies, policy makers, practitioners, researchers, students families, general public	Policymakers and researchers
Stakeholder Engagement	Initially worked with partner agencies to identify questions they wanted the LDS to answer. Marketing plan in place. KYSTATS hosts annual data use conferences and facilitates stakeholder meetings. 'Advisory Committees' made up of partners as well as other stakeholders, including data requesters.	Initially the Center asked stakeholders for research questions with little response. Now the Center takes questions to stakeholders and asks how to tweak to make them more useful. Also performed outreach to researchers and legislators 2 times per year for first 4 years, with sample dashboards and research questions.
Size of Staff	40 people	Approximately 15 FTE (mostly data/IT staff; some admin). Most researchers are affiliated with but not directly employed by the system—about 10 faculty, 10 graduate students.
Data Matching	KCEWS identity resolution and matching process; Developed matching algorithms in-house, through trial and error and based on the matching demographics available for each specific data source.	MLDS performs data matching using personally identifiable information, assigns each record a research identifier, and then removes personally identifiable information.
Data Request Process	Process in place that allows any agency whose data is included in the request to review prior to fulfillment; agency can also review the product (dataset or analysis) prior to it being sent to the requestor.	Approval is majority rules, nobody has veto and agencies cannot opt out if group approves.
Data Request Volume	In 2019: 239 data requests, 578 presentations to stakeholder meetings, 61 published reports	About 10 per year and growing. This does not include reports requested by the legislature.

	Kentucky	Maryland
Research Review Process	Yes, an agency has 10 days to look over product before release.	Yes, 45-day notice period written into MOA. Online description more focused on suppression/ confidentiality of data.
Trust	Board appoints people to help with two-year research agenda (now focused on equity). Update board quarterly on progress.	Continuity of staff at the Center as well as within the agencies.
Sustainability	Currently about 80% grant funded. They have requested state funds in the next biennium budget to cover operations. Key: outreach and stakeholder engagement.	Total appropriation is about $2 million. This covers some admin but mostly data/IT staff.
Key Takeaways for NC	Robust marketing plan- someone meets with every legislator to answer questions about data availability, which has been critical for garnering support. They also host data use conferences with presenters who have used data; legislators are invited.	Center might be closest example we have of an operational data "Switzerland."

	Minnesota	Rhode Island
LDS Name	**SLEDS:** Statewide Longitudinal Education Data System **ECLDS**: Early Childhood Longitudinal Data System	**DataHUB** Rhode Island Longitudinal State System
Established by	Federal SLDS Grant	Federal SLDS Grant
LDS Statute	Minn. Stat. 127A.70	RI H.B. 212
Structure	Centralized	Centralized
Governance	MN P-20 Education Partnership (since 2010 for SLEDS, since 2019 for ECLDS) SLEDS Executive Committee: Commissioners of Office of Higher Ed (OHE), Dept. of Employment and Economic Development (DEED), and Dept. of Ed (MDE) SLEDS Governance Committee: representatives from data contributors and other stakeholder agencies and entities SLEDS Research and Data Advisory Committee: representatives from data contributors and other stakeholder agencies	Rhode Island P-20 Council, an intergovernmental committee comprised of state agents, state representatives, and presidents of RI's public higher education institutions.
Organizational Home	Data managed by SLEDS/ECLDS Coordinators in OHE, DEED, and MDE.	DataHUB maintained by DataSpark, housed at University of Rhode Island. Previously maintained by a nonprofit, which disbanded.
Data Providers	MN Department of Corrections MN Department of Education MN Department of Employment and Economic Development MN Department of Health MN Department of Human Services MN Office of Higher Education	RI Department of Education RI Office of the Postsecondary Commissioner RI Department of Labor and Training RI Department of Children, Youth, and Families DataHub says 20 data sources but doesn't list all.
Independent Colleges and Universities	Initially voluntary, now mandatory. Required to share data in order to receive financial aid. A state law was passed so individual institutions are not reported by name without authorization.	Voluntary. No participation by independents. Three public universities participate because they are required.
Vision Statement	Minnesota has developed the Minnesota Statewide Longitudinal Education Data System (SLEDS) matching student data from pre-kindergarten through completion of postsecondary education and into the workforce. By bridging existing data with other incoming data, a range of education programmatic and delivery questions can be answered to gauge the effectiveness of current programs and design targeted improvement strategies to help students.	The RI DataHUB uses the power of linked data to inform and inspire innovative decision making and research.

	Minnesota	Rhode Island
Example Use Cases	What do we know about the children participating in Minnesota's public early care and education programs (i.e. demographics, disability program combinations)? What is the status of children after participating in the public early care and education programs (i.e. special education use, attendance in the early grades, third grade test scores)?	What is the employment landscape for health field graduates from RI's public higher education system? What do our higher education and workforce data show about RI's production of skilled workers, and whether they are filling high-demand occupations? How do the fields of study of recent postsecondary graduates relate to their place in the state's employment landscape?
Research Agenda	Yes: the 4Ps (Pathways, Progress, Predictors, Performance), which came from stakeholder questions in 2010-2012.	Used to have one; no longer current. Only priority now: whether the request will be useful to Rhode Islanders.
Leadership		
Primary Stakeholders	Research and policy audience; MN legislature.	Primarily state agencies and nonprofits that work with them. Policymakers use one-page bulleted reports.
Stakeholder Engagement	SLEDS partners created a network of regional data coaches to conduct local needs assessments and train stakeholders to use information from SLEDS to answer policy questions. Meet once per year with colleges and universities to court them actively. ECLDS systematically identified stakeholders because early childhood spans HHS, Education, Head Start, so the team had to consider people across all sectors.	Easy to engage agencies, as it was their data. Districts use data insofar as LDS data is incorporated into existing reports.
Size of Staff	7 SLEDS/ECLDS coordinators	6-7 people actively working on DataHUB. DataSpark comprised of 10 people total
Data Matching	SLEDS/ECLDS team identifies unique individuals based on set rules using a probabilistic matching algorithm.	
Data Request Process	Data is retrospective and loaded twice per year (June and November); 5 data access levels established, and a clear data access process in place. Requests are scored against a rubric.	Process is more or less strict depending on risk.
Data Request Volume	As of December 2019: 83 in total (not including data available to colleges and high schools via pre-packaged data mart with select variables): majority from agencies	Substantial volume; requests are not necessarily for integrated data (e.g., 5 currently outstanding researcher requests).
Research Review Process	All requesters must provide SLEDS Coordinator with a paper and electronic report of final results no more than 30 days prior to publication of the study.	
Trust	Long (5-year) process to build trust, particularly between K-12 and Higher Ed: required some staff changes. Early ed was particularly wary about data sharing.	Trust was the main challenge in implementation; ensuring there is equity among participating agencies.

	Minnesota	Rhode Island
Sustainability	State funding for SLEDS from 2014/15 and for ECLDS in 2019. Annual budget between SLEDS and ECLDS is $2 million. Still operating under 2015 grant ($7 million), with supplemental grant received Fall 2019 ($225,000 over 5 years to cover expansion projects). Key: Legislative support, strong administrative processes, staffing.	Funded through grants: Workforce Data Quality Initiative (WDQI) grant, state Department of Health, some private funders.
Key Takeaways for NC	Staffing and IT for the SLEDS is state funded (as of 2014/15), and for ECLDS system (as of 2019). They also fund training and outreach contracts under the state budget. When they bring on a new partner, they train them on how to use the system before accepting them, new partners have to identify how participation will benefit them (as a way to start building sustainability because if they're not using it consistently, they're less likely to be engaged).	Whole system is in flux. Big takeaways from their experience were leadership, stability, neutrality, and capacity, which are key. Had all that prior to the scandal at the nonprofit that was managing the work.

	Washington	Wisconsin
LDS Name	**ERDC** Washington P-20W Statewide Longitudinal Data System	**WISE** Wisconsin Information System for Education
Established by	Statute/Legislation	Statute/Legislation
LDS Statute	Wash. Rev. Code § 43.41.400	Wis. Stat. § 115.297
Structure	Centralized	Federated
Governance	Three committee structure: Data Stewards Committee: comprised of staff and researchers from data contributors and ERDC; ensure the data is understood and used correctly Research and Reporting Coordination Committee: comprised of ERDC staff and representatives from agencies and stakeholder organizations; make recommendations on research priorities) Data Custodians Committee: comprised of technical experts from contributing organizations and ERDC; ensures data is delivered and protected The Legislative Evaluation and Accountability Program Committee also conducts analyses of programs across P-20 and the workforce.	Lightweight governance structure intended to complement and facilitate each agency's existing internal procedures. Chaired by State Superintendent Cabinet; Wisconsin LDS Executive Steering Committee is composed of WDPI staff.
Organizational Home	Data maintained by the Education Research and Data Center (ERDC) under the Office of Financial Management (OFM).	Data managed by WDPI.
Data Providers	Independent Colleges of WA WA Council of Presidents (public 4-year IHEs) WA Employment Security Department WA State Board for Community & Technical Colleges WA State Department of Early Learning WA State Office of Superintendent of Public Instruction WA Student Achievement Council (financial aid) WA Workforce Training and Education Coordinating Board	University of WI System WI Association of Independent Colleges and Universities WI Department of Public Instruction WI Technical College System Workforce data not currently included.
Independent Colleges and Universities	Mandatory as of 2019. Required to share data in order to continue receiving financial aid, however record-level data is only to be provided in response to specific requests (not on a recurring basis).	Voluntary. ICUs can participate, but are treated as a pseudo system; WISE has not built out the functionality yet.

	Washington	Wisconsin
Vision Statement	The purpose of the Washington P-20 and workforce data system is to increase understanding of educational programs and outcomes and relationships to employment programs and outcomes. The system links education and workforce sectors and focuses on the transitions between these sectors. It provides unique opportunities for collaboration among the partner agencies to work together.	The SLDS Project is intended to create tools to facilitate data-driven decision-making for school and district improvement, and to assist educators looking to raise individual student achievement and close achievement gaps.
Example Use Cases	How do the performance profiles of high mobility students compare to those of other students, e.g., attendance, proficiency, graduation, and post-secondary enrollment? What are the most common characteristics of the teacher workforce in schools that show the greatest success with students? What are the demographic, mobility, program, class, grade, and course-taking profiles of students who do and do not achieve and what are their outcomes?	How does Wisconsin Shares funding correlate with early childhood suspensions? (DCF and DPI data) Are there differences in 3rd grade reading proficiency for children who received Wisconsin Shares and were in center-based versus family child care? (DCF and DPI data) What is the behavioral health impact of hearing loss among school age children who participated in birth to 3 special education services? (DHS and DPI data)
Research Agenda	Yes, under "Research Priorities" on website; primarily used to motivate stakeholders consider what is important to them. Categories are: student profile, quality/ achievement, transition/ advancement outcomes, program effectiveness & costs, and teachers.	Yes, set by 3 contributing agencies. Categories: contextual data for equity gaps; access to high quality educational opportunities; access to key educational resources; college and career readiness; early childhood education; social and emotional learning; advancing equity through data reporting.
Leadership	Governor Christine Gregoire (2005–2013) convened the Washington Learns blue ribbon commission in 2005 to review the state's education system and make recommendations for improvements across the education continuum; the Education Research and Data Center (ERDC) was established in 2007 via legislation.	
Primary Stakeholders	Partnering agencies, administrators and legislators, students, parents, teachers.	Districts and schools are the main stakeholders, all agencies involved, researchers, legislators.
Stakeholder Engagement	Half of stakeholders had existing relationships with EDRC, so the EDRC simply had to leverage those relationships. Legislation required participation otherwise.	
Size of Staff	8 FTE	

	Washington	Wisconsin
Data Matching	ERDC does the matching and assigns a research ID using personally identifiable data outside the data system; match depends on source data, as each source has different elements.	Matching engine performs matching by algorithm (via name, birth date, etc.) automatically to relevant data fed into it. Resulting dataset contains a unique research ID with no other agency-level personal identifiers.
Data Request Process	Process in place. ERDC can supply aggregate or de-identified data without approval panel IF project falls within pre-authorized critical questions. Otherwise, goes through a Data Request Panel composed of data contributors.	Data requests are restricted to research questions set by the three agencies; there is a policy in place that the person doing the packaging/extracting of the data should not be the person performing the analysis, so that there is one more step between the people working with the PII and the people working on the research with the program data.
Data Request Volume	63 requests in 2019, 5 of which were denied. Requests are primarily from agencies. Funded studies from legislature not included in count.	1-2 requests per year
Research Review Process	The ERDC confers with data contributors before responding to requests, and requestors must provide ERDC/data contributors with research design and drafts of findings.	Product is re-routed back to governance before publication to be verified/reviewed by data contributors; this is time limited.
Trust	Less important to establish due to statute. Legislation requires all state agencies involved in education, as well as public postsecondary institutions, to create data-sharing agreements with the ERDC.	Being able to work across political change, WISE team spent a lot of time showing the value of what they are able to do. The importance of finding out what is timely, relevant, and actionable; they have a diagram of the "mine, yours and ours." Sweet spot is "ours"... policy relevant, actionable research.
Sustainability	Sustained through grants: ARRA ($17.3 million) used to build data warehouse; currently still using 2015 WDQI and SLDS grants. Key: Legislative champions	Each state agency contributes funds and has absorbed the cost of the data-matching engine; also receive grants. Key: effective team and systematizing processes, as well as continued stakeholder engagement.
Key Takeaways for NC	State statute requires participation by data contributors, and the system has been aided by legislative champions over the years. Legislators are also prime users of the data produced.	There are two systems: ECIDS (early childhood) and WISE (the SLDS); federated model suits those involved. Divided political power in WI has made some things harder, but there has been common ground between parties on data-driven decision-making.

Appendix 7: Resource guide

Actionable Intelligence for Social Policy

- Case studies
- Centering Racial Equity Throughout Data Integration
- Community Engagement Toolkit
- Presentations
- Reports

Data Quality Campaign

- State examples of LDS implementation
- Use cases
- Research partnerships using LDS data

Systemwide Longitudinal Data Systems Grant Program

- Best Practices
- Common Data Definitions
- Guides
- Issue Briefs
- State Spotlights
- Target Team Publications