



The North Carolina Longitudinal Data System: An Update on Progress
Annual Report to the N.C. Joint Legislative Education Oversight Committee, N.C.
Joint Legislative Commission on Government Operations, and N.C. Joint Legislative
Oversight Committee on Information Technology

July 1, 2023

James A. Weaver
State Chief Information Officer and Secretary
N.C. Department of Information Technology

Contents

Executive Summary 3

Background 4

 Purpose of this Report..... 4

 About the North Carolina Longitudinal Data System 4

Update on 2022-23 Activities 6

 Enactment of the Governance Structure of NCLDS..... 6

 Memorandum of Understanding..... 6

 Administrative Code 6

 Governor’s Executive Order 7

 NCLDS Advisory Board..... 7

 Strategic Plan for Implementation of Services 7

 Progress on Governance, Data Quality, and Data Availability 10

 Governance Subcommittees 10

 Data Quality Subcommittees..... 10

 Planned Availability of NCLDS Data, Products, and Services..... 10

 Progress on and Lessons Learned from Use Cases..... 11

Priorities for 2023-24 12

 Automation of the Data Request Process..... 12

 Availability of Publicly Accessible Tools and Products..... 12

 Preparation for Expansion of Available Data 12

Executive Summary

This report provides an update on the development of the North Carolina Longitudinal Data System (NCLDS) and summarizes the ongoing expansion of that system.

The concept of a state longitudinal data system was first approved by the N.C. General Assembly in 2012. The purpose of NCLDS is to link record-level data across state agencies and institutions, support data exchange among those agencies and institutions, and produce information that can be used to guide state decision-making.

Since January 2021, the NCLDS team has made progress in the establishment of a governance structure, development of a guiding strategic plan, and execution of the steps outlined in the strategic plan that will result in a responsive and responsible data system. In 2023, work has included overseeing use cases that test both current and planned components of the system; engaging with potential users of the system (including practitioners, state and local government officials, policymakers, researchers and evaluators); and developing and beginning to execute advanced technical specifications for the system.

For the remainder of 2023 and into 2024, the focus will be on moving from test cases and concepts into full production of NCLDS' public and restricted-access products and services. In addition, the NCLDS team will begin to lay the groundwork for supporting an enterprise framework that allows for the integration of data from a broader array of state agencies and other partners.

Background

Purpose of this Report

In accordance with [N.C.G.S. 116E-4\(c\)](#), this report provides an update on North Carolina Longitudinal Data System (NCLDS, or “the system”) activities, along with information about proposed and planned expansions of NCLDS data. The report concludes with tasks that the N.C. Department of Information Technology’s Government Data Analytics Center will undertake in 2023-24 to move NCLDS closer to full functionality.

About the North Carolina Longitudinal Data System

The concept of a state longitudinal data system was first approved by the N.C. General Assembly in 2012, with significant legislative amendments in 2013, 2016, and 2019 ([N.C.G.S. 116E](#)). Per statute, NCLDS is “a statewide data system that contains individual-level student data and workforce data from all levels of education and the State’s workforce.” The purpose of NCLDS is to “(1) Facilitate and enable the exchange of student data among agencies and institutions within the State; (2) Generate timely and accurate information about student performance that can be used to improve the State’s education system and guide decision makers at all levels; [and] (3) Facilitate and enable the linkage of student data and workforce data.”

NCLDS is led by an executive director and is governed by an advisory board comprising representatives from each state agency that contributes data (“sector contributors,” which currently include the N.C. Department of Health and Human Services, N.C. Department of Public Instruction, N.C. Community College System, University of North Carolina System, and N.C. Department of Commerce) and a representative from the N.C. Independent Colleges and Universities, which also contributes data.

Per the enabling statutory language, NCLDS is not a passive data storage utility; instead, its mission is to provide sector contributors and other trusted practitioner and research partners with secure access to cross-sector, longitudinal, and linked record-level data in order to help them address the state’s most pressing questions via data-informed decision-making, continuous improvement, and performance management processes. More specifically, NCLDS is:

- A data exchange and research **strategic partnership** of the sector contributors listed above
- A **process** for linking, de-identifying, and packaging cross-sector data from participating longitudinal databases (currently the Early Childhood Integrated Data System, NC SchoolWorks, and Common Follow-Up System)

2023 NCLDS Report: Update on Progress

- A **tool** for securely delivering those data to partners, other state entities, researchers, and practitioners for research, evaluation, decision-making, and reporting
- A **common space** for coordinating cross-sector, longitudinal learning goals for the state

When fully operational, the NCLDS team will:

- Oversee implementation of sector contributor-vetted governance procedures for NCLDS-accessible data use
- Produce secure longitudinal record-level and aggregated data packages and products in response to sector contributors' needs and sector contributor-reviewed external requests
- Produce aggregated and deidentified longitudinal data products for public use

NCLDS is within the N.C. Department of Information Technology's Government Data Analytics Center, which works with state agencies to define their analytic needs, determine access to required data sources, and provide staffing in support of NCLDS product development and services.

Update on 2022-23 Activities

Accomplishments since July 2022 fall into two categories: *Enactment of the Governance Structure of NCLDS* and *Progress on Governance, Data Quality, and Data Availability*.

Enactment of the Governance Structure of NCLDS

In 2019, the N.C. Government Data Analytics Center worked with the University of North Carolina at Chapel Hill to develop a strategic plan that would align NCLDS with the needs and priorities of policymakers and sector contributors. North Carolina stakeholders, experts from other states, and evidence-based best practices shaped the contents of the report, which evolved into a roadmap for modernizing NCLDS. Four major structural achievements followed the completion of that report between January 2021 and July 2022, each of which clarified foundational aspects of the emerging system.

Memorandum of Understanding

The foundational document for the NCLDS structure—a Memorandum of Understanding for the Exchange of Longitudinal Data Between Contributors—was executed in June 2020, with the support of the N.C. secretary of state, state superintendent of public instruction, secretary of health and human services, and president of the community college system. The memorandum was expanded in early 2021 to include support from the UNC System, N.C. Department of Commerce, and N.C. Independent Colleges and Universities.

In addition to clarifying key definitions and noting applicable data privacy laws, the memorandum also outlines sector contributor agreements for reviewing and approving data requests, as well as rules for transferring data into and accessing data in the system, exchanging data among signatories, and disclosing data to approved third-party requestors. Also included are auditing and tracking rules, procedures for handling security incidents, and general liabilities.

Administrative Code

Effective January 2021, [Administrative Code 09 NCAC 06D](#) outlines rules governing the NCLDS data request process, requirements for data-sharing agreements between contributors and the N.C. Government Data Analytics Center, and requirements for data-sharing agreements between requestors (both contributor requestors and third-party requestors) and NCLDS contributor owners of the data being requested.

Governor's Executive Order

Governor Cooper's [Executive Order 249](#) (February 2022) further formalized the governance structure for NCLDS by detailing membership of the governance board, establishing the first two committees of that board (the Data Governance and Data Quality committees), and outlining board responsibilities.

NCLDS Advisory Board

Current members of the advisory board include:

- N.C. Department of Health and Human Services: Chief Data Officer Jessie Tenenbaum, as designated by Secretary Kody Kinsley
- N.C. Department of Public Instruction: Superintendent Catherine Truitt (board vice-chair)
- Community College System: TBD
- University of North Carolina System: Senior Vice President for Strategy and Policy Andrew Kelly, as designated by President Peter Hans
- N.C. Department of Commerce: Secretary Mabelle Sanders (board chair)
- N.C. Independent Colleges and Universities: President Hope Williams
- N.C. Department of Information Technology: Secretary James Weaver (non-voting)
- Office of the Governor: Governor's Education Advisor Geoff Coltrane (non-voting), as designated by Governor Roy Cooper

The advisory board met twice in 2022-23 (November 2022 and May 2023), with formal transition of chair responsibilities from Commerce Secretary Mabelle Sanders to Superintendent of Public Instruction Catherine Truitt (per Executive Order 249) taking place at the conclusion of the May 2023 meeting. Andrew Kelly, University of North Carolina System Office senior vice president for strategy and policy, was named vice-chair for 2023-24.

Strategic Plan for Implementation of Services

As of this year, NCLDS and the staff who support it are able to *manually* complete all of the system's basic initial functions:

- Receiving variable-level data requests
- Guiding requests through a review process that involves all relevant data owners
- Completing all necessary legal agreements
- Assembling data listed in an approved request
- Matching data across data sources via application of available identification elements and methods
- Packaging and securely transmitting data

The process is not yet reflective of the efficiency that will be gained once some of these functions are transitioned to a system that either completes them automatically or better supports their manual completion. The transition from the current manual approach to data management and assembly to a more automated approach is the focus of the executive director's **four-phase NCLDS strategic plan**.

Completion of Phase 1. The NCLDS team completed all critical components of Phase 1 by the end of 2022, including:

- Establishment of boards and committees
 - Membership for the Advisory Board and Data Governance and Data Quality Committees has been named, and all bodies have convened multiple times.
- Development of a strategic plan
 - The initial structure and product parameters of NCLDS have been defined.
 - NCLDS goals, target short- and mid-term outcomes, and a phased opening plan have been established.
- Architecture design
 - The fundamental data-related architectural aspects of NCLDS have been established.
- Oversight of governance and data quality
 - All available analysis capabilities have been assessed.
 - External analysis partner recruitment has been initiated.
 - Initial analysis areas of focus (learning goals) have been established.
- Product and services availability
 - A full list of projected NCLDS products has been developed.
 - *Service offered:* On-request dataset creation.

Progress on Phase 2. With input from the advisory board, its two formal committees (Data Governance and Data Quality), and a steering workgroup, the NCLDS team is on target to complete all Phase 2 tasks by the end of 2023. These tasks include:

- Establishment of boards and committees
 - Determination of an oversight process for NCLDS data not managed by one of the current contributors
- Development of a strategic plan
 - Identification of longer-term NCLDS organizational outcomes
 - Completion of a full strategic plan
- Architecture design
 - Critical review of all data-matching procedures and resolution of key cross-sector data-matching challenges

2023 NCLDS Report: Update on Progress

- Establishment of technical requirements for NCLDS tools (data selector/project management user interface/tool, etc.)
- Resolution of data movement security requirements
- Oversight of governance and data quality
 - Development of a dynamic data dictionary and standardized versions of common elements
 - Development of standard NCLDS request and review processes
 - Determination of system design needs for each NCLDS user audience
 - Identification of desired public aggregated data tools
- Product and services availability
 - Establishment of baseline data privacy standards
 - *Products available*: strategic plan; learning goals; data dictionary (beta)
 - *Service offered*: Process for sharing data among contributors
- Outreach expansion
 - Development of a public website and a schedule for regularly updated NCLDS data products

Projected Activities in Phases 3 and 4. Work on the full roll-out of all NCLDS products, services, and supporting governance and quality standards will continue next year in phases 3 and 4. Highlights will include:

- Architecture design
 - Completion of a data selector and project management tool
 - Refinement and improvement of matching procedures and tools
- Oversight of governance and data quality
 - Identification of data elements targeted for cross-sector consistency in definition and quality
 - Resolution of remaining known major data quality issues
 - Refinement of system design needs for each NCLDS user audience
 - Development of a data access handbook
 - Creation of a plan for expanding the data elements available to NCLDS
 - Completion of data management standards (privacy, security, archiving, ethics, and training)
 - Development of a long-term vision for research and evaluation support
- Product and services availability
 - Refinement of data privacy standards
 - Establishment of a pre-publication product review process
 - Establishment of publication quality standards
 - *Services*: Technology-supported data "fetching" and linking; data validation/quality assurance
 - *Products*: Data selector and project management tool; research-ready datasets; data portals for practitioners; goals-aligned research and evaluation projects; public aggregated data products

2023 NCLDS Report: Update on Progress

- Outreach expansion
 - Pursuit of external funding support

Progress on Governance, Data Quality, and Data Availability

As of this writing, the Data Governance and Data Quality committees are in the process of addressing several of the strategic plan elements targeted for completion this calendar year (see *Progress on Phase 2*, above). Each committee's work is supported by several subcommittees:

Governance Subcommittees

- Data availability and access
- Data request review
- Privacy, security, and ethical use of data

Data Quality Subcommittees

- Data lexicon
- Common data elements
- Data matching

Planned Availability of NCLDS Data, Products, and Services

As noted above, the strategic plan includes processes for developing several key NCLDS products and services:

Development of NCLDS' Major Products. NCLDS can manually create on-request datasets for specific projects. Work is underway now to automate some aspects of the on-request dataset creation process and to develop two other data products:

- *Data portals for practitioners*, which are access-protected, cross-sector data products tailor-made by NCLDS for practitioner audiences
- *Research-ready datasets*, which are pre-prepared and pre-vetted (i.e., their content will not require additional review when requested), de-identified, linkable datasets comprising each contributor's most commonly requested data elements

Development of Request Tools. Technical specifications for NCLDS tools designed to automate the request process have been developed and shared with NCLDS' external programming partner with anticipated delivery of development timelines for these tools set for the fourth quarter of 2023.

Incorporation of New Data. In preparation for addressing the strategic plan component related to expanding the data elements available to NCLDS, the team

2023 NCLDS Report: Update on Progress

has started onboarding data elements that are not available to the system via its current partners but will be needed to more thoroughly address questions about workforce preparation. This initial onboarding includes data from state agencies that issue non-degree credentials and licenses.

Progress on and Lessons Learned from Use Cases

To test the viability of the system as outlined by the memorandum of understanding and Administrative Code 09 NCAC 06D and to demonstrate the capabilities of the system, NCLDS has sponsored several use cases since 2021. Each use case involves data from at least two contributors, and several also include third-party partnerships.

Collectively, the use cases incorporate data from all sector contributors. Topics include connections between high school, postsecondary, and workforce outcomes; connections between early childhood and early-grade outcomes; and connections between participation in College and Career Promise programs and workforce outcomes.

The first use case (high school-to-postsecondary/workforce connections) has been completed (read the [use case results and summary](#) from the N.C. Department of Commerce). The other two uses cases are still underway, but collectively, all three already have provided critical insights for the development of NCLDS. In particular, the use cases have clarified the need to develop a common, cross-sector Data Use Agreement to reduce the time needed to secure approval for the release of requested data, as well as additional specifications for NCLDS' data dictionary, data selector, and data request process.

Priorities for 2023-24

Automation of the Data Request Process

The primary NCLDS goal for 2023-24 will be to develop the tools that will support automation of many of the data request fulfillment processes that currently must be completed manually, such as selection of data elements, completion of a data request, review of that request, retrieval of requested data, linkage of requested data, preparation of a data package, and secure transmission of that package.

Availability of Publicly Accessible Tools and Products

Other priorities for the coming year include development of the first set of NCLDS products and services that will be made available to the public, including a public-facing website with information about NCLDS data availability and request procedures, as well as initial publicly available aggregated data products (e.g., dashboards, reports, data stories, knowledge visualizations). In addition, the NCLDS team plans to produce test versions of secure, limited-access products (e.g., research-ready datasets and data portals for practitioners).

Preparation for Expansion of Available Data

Current work is being completed with an eye toward eventual expansion of the data available to the system's users. Many of the topics and questions that comprise NCLDS' cross-sector areas of focus can be fully addressed only once NCLDS is able to incorporate data from state agencies and other partners outside the education sector. As noted above and in a previous report, the initial partnership reflects this understanding via its inclusion of both early childhood and workforce data, and NCLDS has already started bringing in data from sources beyond the original partnership. Starting in 2024, NCLDS will begin laying the groundwork for more fully embracing an enterprise mindset that includes data from an even broader spectrum of sources.