V1.4 (July 24, 2025)

NCLDS Data Privacy and Security Standards and Policies

Overview

NCLDS's approach to privacy and security is guided by an overarching philosophy of **accessing the minimum amount of data necessary to address questions of greatest importance to the state**. NCLDS's data management protocols and processes provide guardrails for five connected facets of data privacy and security: **Privacy Standards**, **Data Transfer**, **Data Storage**, **Data Destruction**, and **Data Archiving**.

I. Data Privacy Standards

Policy and Policy Context

NCLDS governance reflects all applicable state and federal laws, regulations, and policies that safeguard the privacy of an individual's personal information. These laws include:

- Federal
  - The federal Privacy Act
  - The federal Family Educational Rights and Privacy Act (FERPA)
  - The federal Individuals with Disabilities education Act (IDEA)
  - The federal Health Insurance Portability and Accountability Act (HIPAA)
  - The federal Criminal Justice Information System (CJIS)
  - The federal Internal Revenue Code (IRC)
  - The Fair Information Practice Principles (FIPPs)
- State
  - The state Identity Theft Protection Act
  - The state definition of Identifying Information
  - The state Data Classification and Handling Policy
  - The state Information Security Manual and Policies
  - Other applicable state privacy policies and laws

The data to which NCLDS has access are provided by several different agencies and organizations. Most data are directly protected by laws and regulations specific to each data source. For instance, K12 student data are protected primarily by FERPA and IDEA, while some non-education early childhood data are protected by HIPAA. In addition, there are privacy and security regulations that apply to all data.

Requests for data from NCLDS can only be made by vetted Requesters who have completed a formal Data Request form, whose request and reasons for requesting the data have been reviewed and validated by *all* of the Contributors that manage the data being requested, and whose request is permitted by applicable law, regulations, and policies

Protection of Personally Identifiable Information

NCLDS guidelines and policies prohibit personally identifiable information from being made public, in compliance with the state and federal laws cited above. Personally identifiable information is removed through a process called “de-identification” before datasets are provided to external Data Requesters. In fact, many Data Requesters will only see summary-level (not person-level) data.

NCLDS also redacts (removes or blocks) individual data elements not otherwise protected by state or federal laws when it is the standing practice of a Data Contributor to do so.

Identity protections even extend to summary (sometimes called “Aggregated”) data when inclusion of even summary-level information potentially could lead to the identification of individuals. For example, no summary data made available by NCLDS can include data about groups whose numbers are so small (such as program-level data for children in an age group that, in that program, includes only four or five individuals) that making them available could potentially allow someone to identify an individual by using just summary data. This process is called “small-cell suppression.”

Current NCLDS Small-Cell Suppression Threshold Standards

Because some of these standards can be challenging to apply correctly in certain situations, Data Requesters can contact NCLDS with questions about application of the standards. In addition, NCLDS will provide illustrative examples in future versions of this brief. Finally, NCLDS and its Data Contributors will provide relevant feedback to Data Requesters as part of the required final product review detailed in Sections 4.1.6, 4.1.7, and 4.1.10 of the NCLDS Data Use Agreement.

- For all person-level value (count or percentage) data reported:
  - Suppress any ***numerator or count <5***
    - In addition, suppress any numerator when the numerator divided by the denominator is ≥95% or ≤5%
    - In any tables in which a corresponding % value is generated from an already-suppressed count <5, also mask the corresponding % value
  - Mask any cell with a ***value of ≥95% or ≤5%***
  - Present as a range at least ***one additional count or %*** in a distribution that otherwise would reveal that a suppressed count or masked % in another cell is ≤5 or ≤5% (e.g., if the true value of one cell is 4% and all other %s in the table would allow a reader to interpolate that the value of that cell is 4%, the distribution must mask the

4% value ***and*** present another value in the distribution as a range instead of as a number [e.g., 15%-20% instead of 16%]; similarly, if the true value of a count is 4 and the addition of all other counts in the distribution would allow a reader to interpolate that the value of that cell is 4, the report must suppress the 4 count ***and*** present at least one other count in the distribution as a range instead of as a number [e.g., 15-20 instead of 16]).

    - Single tables or groups of related tables that include ***total values directly associated with values for subsets of the same population*** must suppress or mask values in the subset portions of the table(s) that, in combination with other subset values and/or the total values, would reveal suppressed or masked subset values (e.g., of 11 students who passed an exam, 8 were males and 3 were females; the subset value 3 would be suppressed by the first rule above, but the subset value 8 ***also*** must be suppressed because it, in conjunction with the total value of 11, reveals the value of the already-suppressed 3).

- In addition, ***for data for minors or students only***: Suppress any cell with ***count of any size when the denominator <20***
- The preceding suppression and masking rules only apply to non-zero values; values of zero do not need to be suppressed or masked
- Additional suppression standards may apply for specific datasets and situations, at the discretion of the Data Owner(s)

Current NCLDS Redaction Standards

Redaction versus Suppression: NCLDS distinguishes between disclosure avoidance[1] techniques that are applied during the data preparation phase (before data are shared with a Requester) and techniques that are applied before the results of data analysis (by Requesters, Contributors, NCLDS, or any other User of approved data accessed via the NCLDS process) are released or published.

For NCLDS purposes, **redaction** takes place primarily during the data preparation disclosure avoidance process, alongside **anonymization** and **de-identification**.

As required by the NCLDS Data Use Agreement, **suppression** is applied by Requesters during their preparation of the results of their use of approved data that they intend to share publicly. Suppression requirements are outlined in the Data Use Agreement and posted on the NCLDS website.[2] As part of the NCLDS Request Review Process, all Data Contributors are given the opportunity to review a Requester's application of suppression rules before the Requester publishes any results derived from analysis of data sourced via the NCLDS Data Request Process.

[1] NCLDS thanks the US Department of Education for the term "disclosure avoidance," which USED defines here: https://studentprivacy.ed.gov/content/disclosure-avoidance

[2] https://nclds.nc.gov/documents/nclds-data-privacy-and-security-policy/download?attachment

Redaction in NCLDS:

- Definition:

  Redaction is the removal of certain data from a dataset *during the NCLDS dataset preparation stage* to ensure that those data are not exposed to an NCLDS Data Requester or other data system after the data preparation stage. Redaction takes place *before* transmission of the prepared dataset but *after* all other data preparation steps (such as entity resolution) have been completed, to ensure that NCLDS is able to access temporarily all data needed for specific preparation steps.

  Redaction is one of several disclosure avoidance techniques NCLDS can employ, alongside suppression, masking, and other techniques. Redaction can be applied at the table level, the record level, or the element level.

- There are three main redaction actions:
  - Most redaction decisions are made by Contributors on a **request-by-request basis**. Contributors use the existing NCLDS Request Review tools, which include options for Contributors to communicate with NCLDS Administrators and Requesters, to communicate these redaction decisions.
  - In some cases, redaction decisions are applied **universally** to data that are provided by Contributors but that have been derived from a third-party source with its own redaction standards (e.g., National Student Clearinghouse). Information about universal redaction will be included in the NCLDS Data Dictionary, when known.
  - Contributors also may require the application of certain redaction standards to any **outcomes** published or otherwise shared by Requesters. These publication redaction standards will be included in the terms of the NCLDS Data Use Agreement.

- Data Requesters are reminded that **redaction and suppressions are two different processes**, and that, per the terms of the NCLDS Data Use Agreement, Requesters also are required to apply NCLDS suppression rules to all results they intend to publish or otherwise share, regardless of whether the data they received has been redacted in any way.

  Per the terms of the NCLDS Data Request and Review processes, all Contributors must have an opportunity to **review all results** that a Requester intends to publish or otherwise share, ahead of publication or sharing. The intent of this review is not to censor the Requester's results but instead to ensure that all redaction and suppression policies have been followed.

II. Secure Transfer of Data to Requester

Policy and Policy Context

Transfer – Data transfer (or movement of data from its original location to a new location) is handled currently via Secure File Transfer Protocol (sFTP).

See the *Statewide Information Security Policies* System and Communications Protection Policy (SC-8 – Transmission Confidentiality and Integrity; SC-40 – Wireless Link Protection) and Media Protection Policy (MP-5 – Media Transport)

Standards

- Default NCLDS package transfer will be SAS sFTP

III. Secure Storage of Data by Requester

Policy and Policy Context

Storage by Data Requesters: The NCLDS Request Approval Process includes completion of privacy and security checks of a Data Requester's data storage space.

In the unlikely event of a data breach after receipt of Contributor data via NCLDS and before those data have been removed from the Requester's storage device per the data destruction expectations outlined below, a Data Requester must comply with the terms of the Data Requester's Data Use Agreement with respect to discontinuation of use of the requested data and reporting of the breach to NCLDS, as well as to the Requester's Institutional Review Board or equivalent entity. In compliance with the *Statewide Information Security Manual*, NCLDS will report the breach to all Contributors that provided data for the original request, as well as to the NC Department of Information Technology's Enterprise Security & Risk Management Office (ESRMO), and will take action as directed by ESRMO. In addition, depending on the nature and severity of the breach, NCLDS and/or any NCLDS Participating Agency may be required to report the breach to its cognizant state or federal agency. Upon determination of cause, the Data Recipient may be prevented from receiving NCLDS data for a period to be determined by NCLDS and the Contributors affected by the breach.

Storage by NCLDS: Data accessed by NCLDS for the purpose of preparing it to fulfill an approved request are stored temporarily on servers housed in a secured environment that has been validated by the North Carolina Department of Information Technology. There are many check-points—both virtual and physical—that an NCLDS team member must navigate before gaining access to NCLDS's servers and networking equipment. Direct access to data being prepared by NCLDS is restricted to only the NCLDS personnel responsible for the management of the system and the processing of data requests—not even the NCLDS Executive Director has direct access to the data.

In the unlikely event of a breach of the NCLDS environment, NCLDS will suspend all data activities and follow the State's incident reporting policy. NCLDS also will notify all affected Data Contributors within 24 hours, as well as the NC Department of Information Technology's Enterprise Security & Risk Management Office (ESRMO), and will take action as directed by ESRMO. Once the breach has been resolved, NCLDS and affected Data Contributors will notify any affected individuals.

See the *Statewide Information Security Manual* (Sections 3 and 5) and the *Statewide Information Security Policies* Access Control Policy (AC-1—Policy and Procedures), System and Information Integrity Policy (SI-1—Policy and Procedures), and Media Protection Policy (MP-4—Media Storage).

Standards

- NCLDS Contributor MOU, Section 10: IT Security Incidents and Data Breaches
- NCDIT Privacy Threshold Analysis Form
- The National Institute of Standards and Technology (NIST) Risk Management Framework

IV. Requester Data Destruction Expectations

Policy and Policy Context

The legal agreement a Data Requester signs before receiving access to data includes a commitment on the part of the Requester to thoroughly and completely eliminate from its storage facility all data received from NCLDS within 30 days of completion of the project for which the data were requested. The Requester also must verify formally that the destruction has taken place.

See the *Statewide Information Security Policies* Media Protection Policy (MP-6—Media Sanitization [sub-section: Media Disposal]).

Standards

- All NCLDS DUAs will include specific requirements for data destruction
- Default data destruction standard: Immediate and complete destruction upon completion of approved project or end of relevant DUA(s), whichever comes first
- Data destruction expectations will be incorporated into the **NCLDS Project Tracker** (link to be provided after completion of initial working draft of tool), which is a tool for tracking completion of all steps in the NCLDS data request approval and fulfillment process

V. NCLDS Data Archiving

Policy and Policy Context

In some cases, NCLDS will securely archive a snapshot of a dataset prepared in response to an approved Request. These archived copies ensure that a Requester's analyses can be replicated and verified, should questions arise about the validity, reliability, or veracity of any results generated using NCLDS-provided data. Archived data are maintained for 10 years to allow reasonable time for receipt of requests for replication of a given study.

See the *Statewide Information Security Policies*: Media Protection Policy (MP-4—Media Storage [sub-section: Media Archival]).

Standards

- Rationale for archiving: Archiving allows for replication of analyses and verification of analysis results, should questions arise about the validity, reliability, or veracity of any results generated using NCLDS-provided data
- Archived Data
    - Final version of dataset(s) prepared by NCLDS, before transmission to Requester
        - **Current Standards**:
            - Preservation of copy of dataset(s) for **10 years**
            - When an approved dataset that has been archived is formally corrected or updated (e.g., if NCLDS, a Contributor or Contributors, and/or the Requester identify and correct an error in the original approved dataset), NCLDS will destroy the original archived version and create a new archived version based on the revised/corrected dataset
        - *Consideration*: Some prepared NCLDS datasets will include data from otherwise-federated sources. Therefore, archived datasets will not be accessible or usable for any purpose without completion of the NCLDS Data Request process. In addition, since archiving will take place only for datasets requested by 3rd-party Requesters (see below), no archived datasets will include identifiers of any type
        - *Consideration*: Some Contributor data (e.g., wage data) can be viewed only by the Contributor and not by even NCLDS staff; policies will need to be flexible enough to include different archiving procedures for such data
    - Final version(s) of dataset(s) after manipulation by Requester, as used by Requester to generate published outputs
        - **Current Standard**: Archive Requester **formal description of methods** employed for use and analysis of NCLDS data, including such elements as a description of the study sample, data cleaning processes, data inclusion/exclusion rules, data analysis techniques, and summary of results

- Final versions of datasets after manipulation by Contributors for reporting or other purposes not associated with formal research or evaluation studies
  - **Current Standard**: No archiving requirements