GEORGIA'S CROSS-AGENCY CHILD DATA SYSTEM POLICY MANUAL

Table of Contents

i. CACDS Defined 4

ii. Policy Manual Purpose 4

iii. Disclaimer 4

iv. Recommendation for Citation Use 4

v. Feedback and Questions 4

Section 1: Overview of Georgia's Cross-Agency Child Data System (CACDS) 5

a) Introduction to Georgia's Cross-Agency Child Data System (CACDS) 5

b) Vision of CACDS 5

c) Goals of CACDS 5

d) Functions of CACDS 6

Section 2: CACDS Commonly Used Terms and Definitions 7

Section 3: CACDS Governance and Administrative Structure 11

a) Governance Structure 11

i. Children's Cabinet 11

ii. Executive Committee 11

iii. Leadership Committee 11

iv. Research Committee 11

b) Administrative Structure 11

i. CACDS Administering Agency (CAA) 11

ii. Identified Executive Committee Members 12

iii. Identified Leadership Members 12

iv. Lead Identified Researcher and Identified Researchers 12

Section 4: CACDS Data 12

a) Agencies Contributing Data to CACDS 12

b) Programs Contributing Data to CACDS 13

c) Data User Categories 14

d) Available Resources 14

Section 5: Data Levels 16

Section 6: Process for Requesting Data Access 17

a) Research Requests 18

b) Data Requests 23

Section 7: CACDS Technology and Data 24
a) CACDS Technology Overview 24
i. Unique Identifiers 24
ii. CACDS Data Integration and Alignment 24
iii. User Portal 24
iv. Identified Researcher Portal 24
v. Dashboards and Research Reports 24
b) Data Use/Data Ownerships 25
i. Data Privacy and Confidentiality 25
ii. Data Breaches 25
iii. Small Cell Suppression 25

i. CACDS Defined

Georgia's Cross-Agency Child Data System (CACDS) is a longitudinal data system encompassing early childhood data from multiple government-funded programs and organizations serving children and families in the state. Data can be matched across multiple programs, services, and other data sources. The purpose of CACDS is to provide state leaders, policymakers, and other stakeholders with data and reports that can inform decisions, improve programs, and create policies and programs that support children and families. Many states have similar systems broadly referred to as Early Childhood Integrated Data Systems (ECIDS). CACDS is Georgia's ECIDS.

ii. Policy Manual Purpose

The purpose of this Policy Manual is to provide stakeholders and all CACDS users with written guidance on using and accessing CACDS. It is meant to be a living document with revisions made as needed and when new data sources are added. The document is maintained by Bright from the Start: Georgia Department of Early Care and Learning (DECAL) staff, with changes approved by the CACDS Executive Committee. The most up-to-date version of the document is always posted on the CACDS website, www.gacacds.com.

iii. Disclaimer

This Policy Manual provides a detailed overview of CACDS. Much of this manual content has been informed, with the publishers' knowledge and consent, by North Carolina's ECIDS Data Governance Manual and the Georgia Academic and Workforce Analysis and Research Data System (GA-AWARDS) Researcher Use Policy from the Governor's Office of Student Achievement's P-20 Workforce Longitudinal Data System. Other GA-AWARDS documents were also used to develop this guide.

iv. Recommendation for Citation Use

Users of CACDS data and subsequent products (e.g. presentations, white papers, reports, and peer-reviewed journal articles) are encouraged to use the following citation:

Citation Structure

[Chart Title], [Year] by [Geographic Level]. Georgia's Cross Agency Child Data System. [Date Printed]. Retrieved from www.gacads.com.

Citation Example

Conference Data from CACDS, 2018 by Georgia Department of Early Care and Learning. Georgia's Cross Agency Child Data System. May 2018. Retrieved from www.gacads.com.

v. Feedback and Questions

Submit feedback and questions about CACDS and the Policy Manual, as well as requests for support on the use of CACDS to CACDS@decal.ga.gov . Responses will be provided within two business days.

Section 1:Overview of Georgia's Cross-Agency Child Data System (CACDS)

a) Introduction to Georgia's Cross-Agency Child Data System (CACDS)

Georgia's Cross-Agency Child Data System (CACDS) is a longitudinal data system that houses, aligns, and links early childhood-level (birth to age 5) data from multiple government-funded programs serving children and families in the state. By doing so, CACDS serves as a resource for reviewing trends and conducting research related to programs that meet critical needs of Georgia's children and families.

CACDS uses defined rules to assign a unique identifier to data files for tracking children across programs and services, and across time. The data system is compliant with the Federal Education Rights and Privacy Act (FERPA) of 1974, and all state data security policies and statutes. All CACDS data are secured both in transit and in storage. Although CACDS is a standalone early childhood data system, it has the capacity to link the information to Georgia's Academic and Workforce Analysis and Research Data System (GA-AWARDS), also known as the state's P-20 Workforce Longitudinal Data System.

CACDS presents a new and unique opportunity to provide better support to Georgia's children and families. Data can be matched across multiple programs, various services, and other data sources. This integrated data system presents a detailed picture of the distribution of services across the state and gives stakeholders and researchers opportunities to better understand trends related to participation in critical services and programs. Through easy-to-use data tools and analytics, it is possible to understand better how the state is supporting Georgia's youngest learners and their families. This can be especially important for children and families with complex needs.

b) Vision of CACDS

CACDS was created to be an integrated longitudinal data system to facilitate research on and evaluation of the state's early childhood programs. The data it stores include information on early childhood program participation to help identify gaps in services targeting Georgia's youngest learners and their families. This is especially important in understanding access in critical geographic areas and for certain demographic groups. Access to this data will enable programs to provide better services that meet community needs more adequately.

Beyond policy analysis and research, CACDS was designed to improve access to information on early care and learning programs and services by providing data to families, stakeholders, and other entities focused on children and families.

c) Goals of CACDS

The overall goal of CACDS is to serve as a resource to help state leaders, policymakers, program directors, researchers, and other stakeholders ascertain how well Georgia's early childhood programs are meeting the needs of the children and families they serve. Specific goals include:

- Provide critical information, through the CACDS reporting system, about federal and state programs that serve children and their families, including unduplicated counts of those who receive services and participate in the state's early care and education system
- Create a data resource that will help state agency leaders, and academic and policy researchers better understand how well these programs meet the needs of the children and families they were designed to serve
- Facilitate alignment across the state agencies that serve and support children and families with high needs

d) Functions of CACDS

CACDS houses data from multiple programs serving children and families. In all, there are over 400 unique, matched data elements across 12 programs. This information can be used to assess early childhood programs and services to understand better how well the state is meeting the needs of its youngest learners. CACDS can also be used to identify service gaps among various early childhood programs across the state's geographic regions and demographic groups, including providing a picture of local level service usage.

CACDS includes a function for requesting data allowing targeted analysis of a range of programs, policies, and issues by stakeholders and policymakers. For example, by utilizing outcome data in CACDS and GA-AWARDS, research projects can be designed to look for associations between program participation and later outcomes. Section 6 details the process for requesting CACDS data access to conduct research.

Specific functions of CACDS include:

- Collecting and housing longitudinal data across multiple programs and services in one data system;
- Aligning this longitudinal data to provide a snapshot of how individuals are accessing services and programs at the state and community level
- Connecting CACDS data with GA-AWARDS longitudinal data system to help facilitate research related to the impact of early childhood programs and services over time
- Creating publicly available reports that provide information about early childhood services at the state and local level and that can be customized to meet community needs
- Serving as a research and evaluation resource for policymakers, program directors, approved researchers, and child-family advocates needing information to answer specific questions

Section 2:CACDS Commonly Used Terms and Definitions

Whenever used in this Policy Manual, the following terms and definitions, unless otherwise stated, have the following meanings:

1. Aggregate-Level Data: Data that are at a group or cohort level and not at an individual level.

2. CACDS Administering Agency (CAA): The agency responsible for housing and managing the data system. Currently the CAA is Bright from the Start: Georgia Department of Early Care and Learning (DECAL). More information about the CAA can be found in Section 3.

3. CACDS Application: The technology components of the data system, including the web portal, data retrieval tools, and reporting package.

4. CACDS Confidential Data: Confidential information at the individual level from the Participating Data Contributors linked in the CACDS database. This confidential data may include, but are not limited to: name, Social Security number, address, birth date, gender identity, race, ethnicity, or disability status.

5. CACDS-GA-AWARDS Confidential Data: Confidential information at the individual level from the Participating Data Contributors that is linked in the CACDS database and is also linked to GA-AWARDS data. This confidential data may include, but are not limited to: name, education history, Social Security number, address, birth date, gender identity, race, ethnicity, or disability status.

6. Cohort: Group of people used in a study or research who have something in common, such as demographic identifiers or program participation.

7. Confidentiality and Data Usage Agreement: An agreement made between the CAA and non-Participating Data Contributor (PDC) data requestors.

8. Data Sharing Agreement (DSA): An agreement signed between a Participating Data Contributors that specifies the data that will be incorporated into CACDS, as well as the terms of their use.

9. Data User Categories: Groupings developed by the CACDS Executive Committee to establish hierarchies and priorities among the universe of potential CACDS data users. Additional information about each category can be found in Section 4. Detailed processes and procedures for each user category can be found in Sections 10-14 respectively. The Data User Categories are:

    a. Participating Data Contributors (PDCs): include agencies that contribute data to CACDS regularly;
    b. Georgia Affiliated Non-Participating State Entities (GANSEs): include certain government agencies and researchers from University System of Georgia (USG) and Technical College System of Georgia (TCSG) schools;
    c. Third Party Managed Users (TPMUs): include external entities conducting research on behalf of a PDC or GANSE;
    d. Georgia Legislators and Policymakers (GLPMs): limited to members of the Georgia State Assembly (State Senators or Representatives), staff from the Governor's Executive Office, and staff from the Governor's Office of Planning and Budget;

e. Others: include such stakeholders as advocacy organizations, non-profit and non-governmental organizations serving children and families; non-USG and non-TCSG affiliated researchers; and other groups with demonstrated interest in child and family well-being.

10. De-identified: Data that cannot be linked to a named individual. At a minimum, the following fields would be removed to de-identify any data set: first name, last name, middle name or initial, Social Security number, address, birth date, and gender. In some cases, fields such as race, ethnicity, and disability status will also be removed if the cell size is so small that individuals can be easily identified.

11. Executive Committee: The group responsible for CACDS governance and decision-making. The Executive Committee is composed of one senior level official from each Participating Data Contributor (PDC) who is authorized to make decisions on behalf of that agency.

12. External Research and Reporting: Activities utilizing the categories of CACDS data that are accessible to Third Parties as defined in this Policy Manual.

13. Family Educational Rights and Privacy Act (FERPA): A federal law that protects the privacy of student education records.

14. Georgia's Academic and Workforce Analysis and Research Data System (GA-AWARDS): The state's longitudinal research data system that links Pre-K through postsecondary student information and Department of Labor data. Also known as Georgia's P-20 Workforce Longitudinal Data System. CACDS was designed to link other relevant state agency data to GA-AWARDS. Agencies that participate in either CACDS or GA-AWARDS, but not both, must also meet the guidelines of the data system to which they do not contribute data to access a data set that includes both.

15. Georgia Affiliated Non-Participating State Entities (GANSEs): A Data User Category that includes representatives and researchers of USG and TCSG institutions, and state agencies that are not Participating Data Contributors (PDCs) but who may be granted access to certain types of CACDS data and reports for certain types of research.

16. Georgia Legislators and Policymakers (GLPMs): A Data User Category that includes members of the Georgia General Assembly (State Representatives and Senators) and their staffs, the Governor's Executive Office staff, and the Governor's Office of Planning and Budget staff who are granted expedited access to CACDS data and reports.

17. Identified Researcher (IR): A person identified by a Participating Data Contributor (PDC) or identified in an accepted research request as possessing the high-level analytical skills, research training, and credentials needed to mine CACDS data and answer critical educational policy and evaluation questions. Identified researchers have access to de-identified individual-level data to conduct research and prepare reports. There are several types of IRs who may interact with CACDS:
    a. PDC-Identified Researcher: Researchers identified by Participating Data Contributors
    b. Lead PDC-Identified Researcher: The lead researcher identified by each Participating Data Contributor
    c. Non-PDC-Identified Researcher: Researchers identified in data requests or in contracts with a Participating Data Contributors

d. Lead Non-PDC Identified Researcher: The lead researcher identified in data requests or in contacts with a Participating Data Contributors

18. Individual-level data: Information in a data set that is associated with a specific person, subject, or respondent, rendering the person easier to identify.

19. Internal Research and Reporting: Information that is produced by utilizing CACDS data and is available only to Participating Data Contributors (PDCs) and their designated employees, Executive Committee members, agents of Participating Data Contributors (PDCs), and Georgia Legislators and Policymakers (GLPMs).

20. Misuse: Knowingly accessing, releasing, publishing, disclosing, altering, or using data, research, or reports in CACDS without approval or in a manner that is prohibited or inconsistent with the procedures in this Policy Manual. The Executive Committee determines whether the data have been significantly, intentionally, or repeatedly misused, or utilized in a manner that was not intended and decides whether and how to punish the person or entity responsible for the infraction.

21. Memorandum of understanding (MOU) is an agreement between two or more parties outlined in a formal document. It is not a legally binding document but signals the intention of all parties to move forward with a contract.

22. Others: The Data Use Category covering entities that need access to CACDS data and reports to execute their own research or to support or coordinate research projects with a Participating Data Contributor (PDC) or a Third Party Managed User (TPMU).

23. Participating Data Contributors (PDCs): A Data User Category covering agencies that have signed a Memorandum of Understanding (MOU) or Data Sharing Agreement (DSA) to provide data to CACDS regularly. PDCs have access to a full set of CACDS data that includes de-identified data from other agencies and individual data from their own agency. PDCs can use a copy of their own data to match with U.S. Census and U.S. Postal Service data sets or one that has been cleaned. Also referred to as Participating State Agencies.

24. Personally Identifiable Information (PII): Individual-level data that can be used on their own or with other information to identify, contact, or locate a single person.

25. Sufficient Cell Suppression: Refers to the permissible size of publicly reportable data sets. A data cell, table, or text discussion containing information about five or more children can be reported. To avoid privacy violations or disclosure of PII, data sets with files for less than five children that haven't been can't be reported.

26. Third Party Managed Users (TPMUs): Individuals or entities not employed by a Participating Data Contributor (PDC), but under contract to it to conduct research using CACDS data. TPMUs must report to the PDC's Identified Researcher (IR) and follow the procedures in this Policy Manual. The PDC's

status will determine the TPMU's access to the data sets that link CACDS and Georgia's Academic and Workforce Analysis and Research Data System (GA-AWARDS) data.

27. Third Party: Any non-employee individual, group, corporation, or other entity that is associated with a PDC by contract or other arrangement. (see definition of Data Users).

28. Time-Sensitive Priority Requests: Requests for CACDS data made by the staffs of Governor's Executive Office or Office of Planning and Budget, or members of the Georgia General Assembly and their staffs, to be fulfilled within three business days or less, to the best of the IR's ability.

Section 3:CACDS Governance and Administrative Structure

a) Governance Structure

This section describes the four entities that make up the governance and administrative structure of CACDS. Each body plays a role in ensuring that CACDS meets its stated goals.

i. Children's Cabinet

The Children's Cabinet consists of state agency commissioners and executive directors. These members are ambassadors for Georgia's CACDS at the highest levels of state government. Members of the Children's Cabinet who contribute data to CACDS sign DSAs governing their agency's participation and appoint their agency's representative to the CACDS Executive Committee.

ii. Executive Committee

The Executive Committee is a group of representatives appointed to develop, adopt, and monitor the rules and policies for CACDS. Each PDC has one representative on this committee who votes on process and policy matters. These representatives are typically director-level state agency leaders. Additional staff members may attend Executive Committee meetings, but they may not participate in voting. The Executive Committee meets at least quarterly.

iii. Leadership Committee

The Leadership Committee consists of the program and data experts working with CACDS data feeds from within a PDC. These members deliver feedback to their Executive Committee representatives, who will then speak on their behalf at Executive Committee meetings. The Leadership Committee meets twice a year.

iv. Research Committee

The Research Committee is a group of PDC representatives appointed to review all external research requests and make recommendations to the Executive Committee. Each PDC selects an employee who is a research professional to serve as its representative on the Research Committee. The Research Committee meets at least four times a year and may meet more often, depending on the volume of research requests.

The Research Committee is responsible for creating and maintaining a multi-agency research agenda utilizing CACDS data. As part of this agenda, this group will create at least two research papers each year and determine the lead agency for the reports.

b) Administrative Structure

i. CACDS Administering Agency (CAA)

The CACDS Administering Agency (CAA) is responsible for housing the physical data system. Currently, Bright from the Start: Georgia Department of Early Care and Learning (DECAL) serves as the CAA for CACDS and manages its day-to-day operations. The CAA is responsible for creating a management structure and appointing personnel to:

- Manage daily operations

- Coordinate the bi-annual Leadership Committee meetings, and the quarterly Research Committee and Executive Committee meetings
- Manage communication and research requests
- Oversee the CACDS research agenda

The CAA is also the only entity with access to child data containing personally identifiable information (PII) for the purposes of data matching and cleansing, and research report creation. Only the CAA Manager and a few select staff members are granted this level of access.

ii. Identified Executive Committee Members

Identified Executive Committee Members are appointed by the PDCs as their representatives to make decisions on its behalf. The member is required to attend Executive Committee meetings. From an administrative standpoint, the Executive Committee makes decisions that are then implemented by the CAA.

iii. Identified Leadership Members

Identified Leadership Members are appointed by the PDCs as their representatives on the CACDS Leadership Committee to review the CACDS research agenda, provide specific program and IT updates, and make recommendations that are communicated to the Executive Committee for possible implementation by the CAA. The Leadership Committee meets 2 to 3 times a year.

iv. Lead Identified Researcher and Identified Researchers

The Lead Researcher is appointed by the PDC to be the point of contact for the CAA Manager. Each PDC can also appoint 1 to 4 IRs who can directly access de-identified data. Each IR will be required to complete a training course before receiving access to the data. The training will be developed through a collaboration of the CAA and the CACDS Research and Executive Committees. The training will teach participants to access CACDS data and the importance of data security and privacy. Upon completion of the training, the IR will receive a User ID, which is needed to log into the shared data portal.

Georgia Head Start and Early Head Start Grantees will identify 1 Lead IR to serve on the Research Committee and 1 to 4 IRs who can access de-identified data. The Executive Committee Head Start representative will work with the grantees to appoint its Research Committee representative.

Section 4:CACDS Data

a) Agencies Contributing Data to CACDS

Currently, there are five PDCs contributing data to CACDS[1]. These agencies have signed either an MOU or DSA to provide data to CACDS on a regular basis. These are the current PDCs with an executed DSA:

1. Georgia Department of Human Services, Division of Family and Children Services (DFCS)
2. Georgia Head Start and Early Head Start Grantees
    - Baldwin County School District
    - Burke County Board of Education

[1] This list is subject to change as other government agencies are added or removed by the Executive Committee.

- CDI HS/DeKalb And Rockdale Counties, GA
- CDI West Georgia
- Clarke County School District
- Clayton County Community Services Authority, Inc
- Coastal Georgia Area Community Action Authority, Inc
- Coastal Plain Area Economic Opportunity Authority, Inc
- Action Pact
- CSRA Economic Opportunity Authority, Inc.
- Easter Seals North Georgia
- Economic Opportunity for Savannah Chatham County Area
- Emanuel County School District
- Enrichment Services Program, Inc.
- Family Resource Agency of North GA
- Fort Valley State University
- Bright from the Start: Georgia Dept. of Early Care and Learning
- Jasper County Board of Education
- Macon Bibb County Economic Opportunity Council Inc
- McIntosh Trail Early Childhood Development Council Inc
- Middle Georgia Community Action Agency
- Ninth District Opportunity, Inc.
- Putnam County Board of Education
- Randolph County Board of Education
- Resurgent Education and Community Health Service, Inc.
- Southwest Georgia Community Action Council, Inc
- Tallatoona Community Action Partnership, Inc.
- The Sheltering Arms
- Washington County Chapter Negro Business and Professional Women's Club, Inc.
- YMCA/Early Childhood Development Co., LLC

3. Bright from the Start: Georgia Department of Early Care and Learning (DECAL)
4. Georgia Department of Education (GaDOE)
5. Georgia Department of Public Health (GaDPH)

b) Programs Contributing Data to CACDS

The five PDCs listed in Section 4a above represent 10 programs and services serving children and their families:

- Childcare and Parent Services (CAPS) http://caps.decal.ga.gov
- Early Head Start https://eclkc.ohs.acf.hhs.gov/hslc
- Head Start https://eclkc.ohs.acf.hhs.gov/hslc
- Preschool Special Education Individuals with Disabilities Education Act (IDEA Part B) http://www.gadoe.org/Curriculum-Instruction-and-Assessment/Special-Education-Services/Pages/Preschool-Special-Education.aspx
- Georgia's Pre-K www.decal.ga.gov
- Georgia's Pre-K Rising Pre-K Summer Transition Program
- Rising Kindergarten Summer Transition Program http://www.decal.ga.gov/Prek/SummerTransitionProgram.aspx

- Babies Can't Wait (IDEA Part C) https://dph.georgia.gov/Babies-Cant-Wait
- Children 1st http://dph.georgia.gov/children1st
- Georgia Home Visiting www.greatstartgeorgia.com

In addition to data from the above programs, CACDS includes publicly available U.S. Census Bureau and U.S. Postal Service data. Census data provide community demographic information and Postal Service data facilitate address matching and more accurate geo-mapping. CACDS also includes early care and education provider information provided by DECAL, such as child care licensing information and Quality Rated status.

c) Data User Categories

The CACDS Executive Committee has identified specific Data User Categories. These categories were developed to establish hierarchies and set priority levels for accessing data, approving research requests, and utilizing available CACDS supports. These Data User Categories are:

- Participating Data Contributors (PDCs):
  - Members of this group signed either an MOU or DSA specifying when their data will be pushed to CACDS.
- Georgia-Affiliated Non-Participating State Entities (GANSEs):
  - Federal, state, and local government agencies that have a specific data or research need that can be fulfilled by utilizing linked data from multiple agencies. In most cases, data requests for a single agency will go through that agency's request process.
  - Approved Researchers from USG and TCSG. Approved researchers are employees of a USG or TCSG institution who have submitted a data request that has been approved by the Executive Committee. The approved researcher status is only applicable to data requests that have been approved.
- Third Party Managed Users (TPMUs):
  - This group includes external researchers who are performing research studies on behalf of a PDC or a GANSE. Before receiving access to the data, each TPMU must sign a user agreement ensuring the security of the data.
- Georgia Legislators and Policymakers (GLPMs)
  - Members of the Georgia General Assembly and their staffs, and the staffs of the Governor's Executive Office and Office of Planning and Budget.
- Others
  - Potential users from advocacy organizations, non-profit organizations serving children and their families, researchers not affiliated with USG or TCSG, families, programs, and other stakeholders who do not fit into a delineated category and who are not data contributors.

d) Available Resources

Data and tools are already available to the public via the CACDS website. These reports were created with support from the governance and administrative structure and serve as a starting point for anyone interested in using statewide early childhood data.

- Public Reports (Non-Customized)
  These reports are designed and developed by the CAA with input from the Executive and Leadership Committees. These public reports are available via the CACDS website and cannot be manipulated. A function on the CACDS website allows users to request or suggest additional types of Level 1 reports, which will be considered by the Executive Committee.

- Public Reports (Customized)
  These reports are designed and developed by the CAA with input from the Executive and Leadership Committees. These public reports are available via the CACDS website and can be manipulated. Users can customize the reports to focus on specific geographic areas and programs. A function on the CACDS website allows users to request or suggest additional types of Level 2 reports, which will be considered by the Executive Committee.

Section 5:Data Levels

CACDS consists of four data levels. As outlined below, Level 1 is de-identified, individual-level child data. Level 2 is individual-level child data that contain PII. Level 3 data is de-identified individual-level child data from multiple programs. Level 4 data is Level 3 data that include PII.

- Level 1 – Data Sets of De-Identified Individual-Level Child Data from a Single PDC
    - Level 1 data consist of a de-identified, individual-level child data set that is matched with Census and Postal Service data. This data set has all PII removed and only consists of data provided by the PDC to CACDS, plus any matching data from public sites such as Census and Postal Service data.
    - CACDS provides each PDC with a copy of its Level 1 data annually. A PDC's Executive Committee member or Lead Researcher can request that the CAA update its data when necessary.
        - All Head Start grantees contributing data to CACDS will receive a copy of its Level 1 data. The Executive Committee Head Start representative and Head Start Lead Researcher will receive Level 1 data that include all information contributed by Head Start grantees.
    - Requests from a non-PDC for access to a PDC's Level 1 data can be submitted to that PDC and do not require Executive Committee approval. This rule applies to requests related to projects that the PDC has contracted for with a Third Party. The CAA should assist the PDC in fulfilling such requests and in supporting these external contracts.

- Level 2 – Data Sets of Individual-Level Child Data from a Single PDC Containing PII:
    - Level 2 data consist of an individual-level child data set that contains PII and is matched with Census and Postal Service data. The data set contains only data provided to CACDS, plus any matching data from public sites such as the Census and Postal Service.
        - One of the benefits of PDCs having access to Level 2 data is the ability to use a data set that is matched to Census and Postal Service data. For example, a PDC might find a full data set that includes Census and Postal Service data useful in completing administrative tasks.
    - Level 2 data are not automatically provided to a PDC. The PDC's Executive Committee representative or Lead Researcher must request them from the CAA.
        - The Head Start Executive Committee representative or the Lead Researcher must submit requests for Level 2 data on behalf of Head Start grantees.
        - Level 2 data are provided to the PDC only. If an external party or a Third Party contracted by the PDC needs Level 2 data and can't get them from the PDC, the CAA will provide them after receiving the formal request and gaining approval from the Executive and Research Committees.

- Level 3 – Data Sets of De-Identified Individual-Level Child Data from Multiple PDCs:
    - Level 3 data consist of de-identified, individual-level child data set from multiple PDCs. This data set has all PII removed. It is also matched with other publicly available data sets, such as Census and Postal Service data.
        - There are many benefits to Level 3 data, which form the crux of investments in ECDIS such as CACDS. Level 3 data provide opportunities to create reports, conduct research, and inform policy using a myriad of data sources. For example, Level 3 data can be used

to conduct a research study that could, eventually, look at the impact of participation in multiple programs on later school readiness.

- Each PDC has access to Level 3 data. The data set is created once a year and the Executive Committee member and PDC-approved researchers receive a login that gives them access.
  - Head Start-approved researchers and the Head Start Executive Committee representative have access to Level 3 data and work with Head Start grantees to meet their data needs.
- Non-PDCs may request Level 3 data. These research requests include a review by the CACDS Research and Executive Committees. Section 6 explains the research request process and provides instructions for PDCs working with Third Party users.

- Level 4 – Data Sets of Individual-Level Child Data from Multiple PDCs Containing PII:
  - Level 4 data include individual-level child data from multiple PDCs that contain PII. Only the CAA and a few select staff have access to Level 4 data when cleaning and integrating the data sets.
  - Non-PDCs are not eligible to request Level 4 data.

Section 6:Process for Requesting Data Access

CACDS can provide data to fulfill requests from recognized Data User Categories. This section discusses the three types of requests that recognized users can access from CACDS.

- Research Requests
  - Requests can be made for Level 3 data to be used for broader analysis related to relevant state research questions. This analysis may be published and made available to the public. These requests will go through a tiered request process that is explained later in this section.
- Data Requests
  - These requests are made with the intent to inform internal policy and any analysis or findings will not be published widely. Data can be requested for use in data visualization tools, but only if the data will be kept de-identified and aggregated within the tools. Due to the nature of the request, these requests will come from PDCs primarily, but non-PDCs can also request data. The CAA will fulfill these requests and inform the relevant PDCs of report requests that utilize their data.
- High Priority Reports
  - These are customized reports that answer specific research or policy questions. They are prepared in response to requests from Georgia Legislators and Policymakers (GLPMs) — the Governor's Executive Office, the Governor's Office of Planning and Budget, and members and selected staff of the Georgia General Assembly.

a) Research Requests

The formal request process for Level 3 CACDS datasets was developed to ensure data quality and maintain accuracy in reporting across all products developed using CACDS data. The process for submitting research requests is detailed below. In these requests, an identified researcher (IR) from a non-PDC will work with the CAA (CACDS Administering Agency) Manager to complete the request.

Third Party Managed Users that want access to data through their collaboration with a PDC will have to submit additional data agreement forms after the completion of the formal data request process. These additional documents will be provided by the CACDS Research Committee once the additional securities required are identified.

The following provides a list of the general steps in the data request process:

1. Requester completes the data request application found on the CACDS website and in this policy manual. Application is sent via email to the CAA Coordinator. It is expected that the identified researcher (IR) and the CACDS Administering Agency (CAA) Coordinator will be in communication throughout the entire process of completing the application. (Approximately one- two weeks.)
2. The CAA Coordinator will review the application for completeness and submit to the CACDS Research Committee for a thorough review or send back to the IR. (Approximately three to five business days for completion.)
3. The Research Committee will determine if the project is feasible, aligns with overarching goals of CACDS, and if there are any confidentiality or privacy concerns. The Research Committee will provide an approval or denied request recommendation to the Executive Committee. Each participating agency's lead research representative must provide approval for request to move forward in process. (Expected to take no more than three to four months to determine project feasibility from the initial submission time)
4. The Executive Committee will provide the final decision. Each participating agency's executive committee representative must provide approval for request to be accepted. (Approximately two weeks to a month)
5. Requester and the executive committee representative from each agency providing data must sign the Data Usage Agreement before any data can be shared. (Approximately one month)

The request process will begin with an initial request made through the CACDS Request Portal found on the website (www.gacacds.com). Information that must be submitted using this portal includes:

1. Name
2. Organization affiliation
3. Intent of the research and timeline of project
4. Brief description of the project
5. Brief outline of methodology
6. Brief description of what CACDS data would be needed

This initial request will be submitted online, received by the CAA Coordinator via email. The CAA Coordinator can also review these requests through the website using the secure CAA login. The CAA research and data

team will review the request for its feasibility, legal/confidentiality concerns, and its policy relevance to Georgia's CACDS research agenda. The CAA team will use a rubric based on these three aspects to inform its decision.

The CAA team will either approve or deny the request within 20 business days. If the request is approved to move to the next stage of the request process, the requester will use a secure login to submit a full request using the online proposal platform located on the CACDS website.

This request will require more detailed information building upon the initial request. Information required for this proposal includes:

- Identifying information: Name, email, title, address, phone number, CV, education system affiliation, affiliated organization or agency, and if the organization has an IRB approval process
- Has this project or similar project been submitted previously by you or coauthors?
- Project title
- Project description
- Research questions
- Does this proposal align with any of the following goals included in the state strategic plan?
- Describe how the proposal aligns with this goal/these goals?
- Does this proposal answer any questions included in the CACDS Research Agenda?
- Please list the questions and explain how proposal address these questions
- Is this project designed to be used for a program or policy evaluation?
- Purpose of the project
- Intended audience for the project
- Where do you plan to submit your results for publication or presentation (e.g. journals, conference, etc.)
- Project timeline, listed in days, beginning with the receipt of the data if approved (e.g. 30 days – cleaning and analysis, 15 days – writing, etc.) Please indicate if you have specific deadlines (e.g. grant application deadlines, etc.)
    a. Task 1:
    b. Estimated Timeline (days):
- Are you relying on any literature from peer-reviewed journals, reports, previous studies or projects to inform your project? Select all that apply
- If you are using any literature or reports, please provide a list of references used to guide development of the research questions.
- Which type(s) of analysis will this project use?
- Please provide a methodological description of your project, including the propose model. Explain why this model is the best method to answer the proposed questions.
- Describe how you plan to present the results (e.g. school-level averages, regression coefficients, etc.)
- To the best of your knowledge, please indicate which agency provides the requested data to CACDS.
- For which years are you requesting the data?
- Please upload the element justification template listing the variables you are requesting, along with a justification for each variable (e.g. treatment variables, control variables, how variable will be used in the model).
- If you are requesting data for a subset(s) or the statewide population, please indicate the subset(s) below.

- Are you planning to match CACDS data with any other data source?
- Are you or any other person listed in the application receiving personal financial or monetary compensation for this project?
- How many individuals other than yourself, including graduate assistants, will access the data?
- What roles and specific functions will the lead researcher (as named in the CACDS application) perform in the event the request to use CACDS data is approved?
- Will any CACDS data be provided to a third-party entity such as a research institution or think tank, or an employee therein, as part of the study?

Once the requester has submitted the full proposal online, the CAA Coordinator will be alerted via email. The CAA Coordinator will bring the full proposal to the next scheduled Research Committee Meeting. Using a more detailed rubric that expands upon the initial request rubric, the Research Committee will discuss and review the full proposal. The committee will determine whether the request is denied, approved in full to move forward in the process, or approved with conditions.

If the request is ***approved in full***, it will then be reviewed by the Executive Committee which will vote to grant final approval or rejection.

If the request is ***denied*** by the Research Committee or rejected by the Executive Committee during final approval, the proposal cannot move forward in the process, and requester cannot resubmit the same proposal.

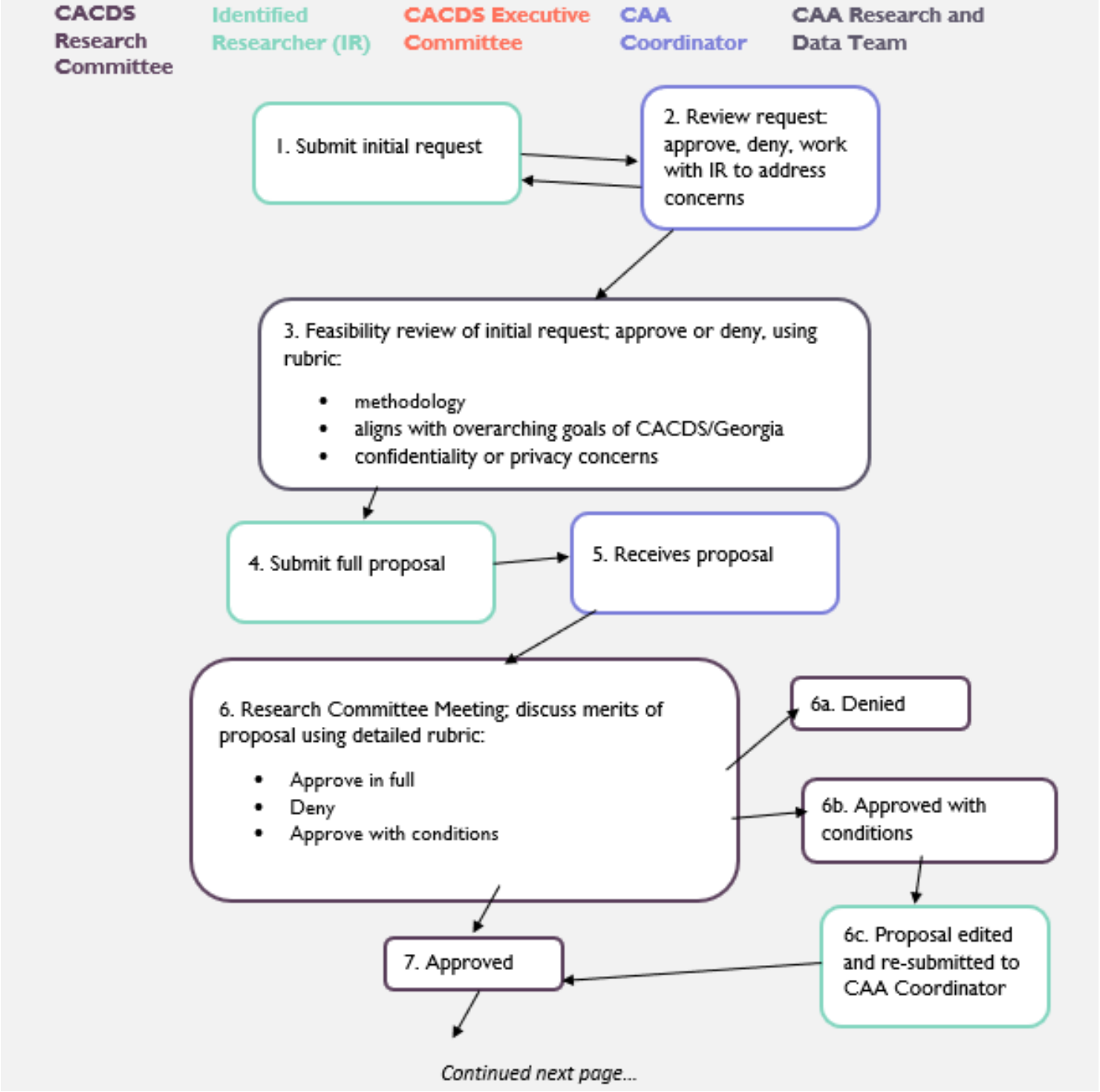
If the request is ***approved with conditions***, the Research Committee will offer edits and suggestions that the requester must address and include for the submission to move forward. Once these edits are added, the CAA Coordinator will verify the changes, and then send the proposal to the Executive Committee for the final vote.

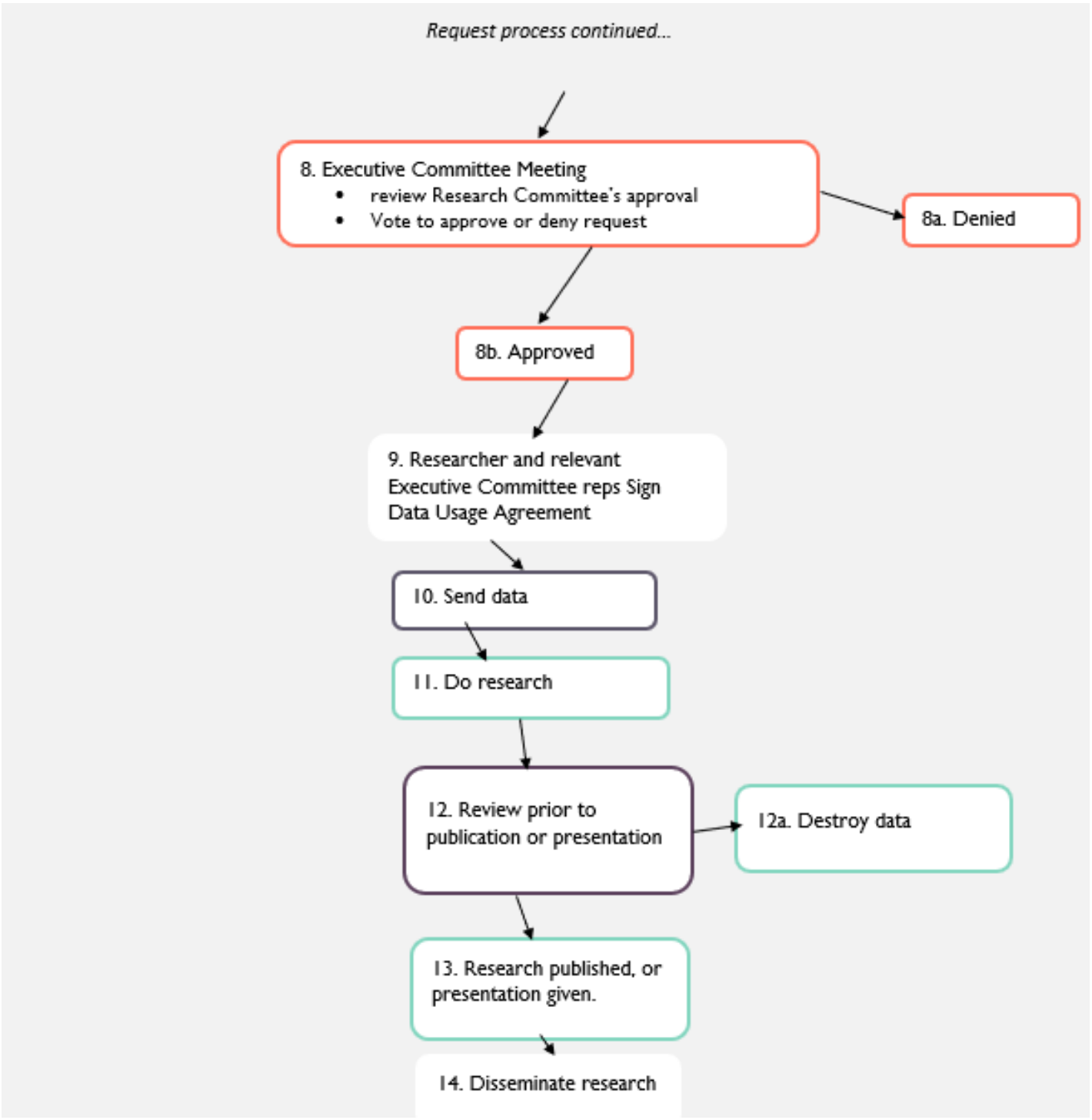
The status of the research request can be monitored via the requester's secure login set up after receiving approval from the CAA. When the request is either approved or denied at any stage of the process, the requester will be alerted via email. After the Executive Committee votes of final approval or rejection of the research request, the CAA will inform the requester of the decision via email and outline the next steps of the process.

If the research request is granted approval, the requester will then be asked to review and sign the Data Usage Agreement. Upon signature, the requester can then follow the steps outlined below by the CAA to retrieve the data securely and for a determined amount of time.

The following flow chart illustrates the full research request process from the initial request to the dissemination of data.[2]

[2]Flow chart inspired by Wisconsin's WISE "Confidential Data Request Process. Request process inspired by Georgia's P-20W integrated data system, GA*AWARDS.


CACDS Research Committee
Identified Researcher (IR)
CACDS Executive Committee
CAA Coordinator
CAA Research and Data Team
1. Submit initial request
2. Review request: approve, deny, work with IR to address concerns
3. Feasibility review of initial request; approve or deny, using rubric:
methodology
aligns with overarching goals of CACDS/Georgia
confidentiality or privacy concerns
4. Submit full proposal
5. Receives proposal
6. Research Committee Meeting; discuss merits of proposal using detailed rubric:
Approve in full
Deny
Approve with conditions
6a. Denied
6b. Approved with conditions
6c. Proposal edited and re-submitted to CAA Coordinator
7. Approved
Continued next page...


Request process continued...
8. Executive Committee Meeting
review Research Committee's approval
Vote to approve or deny request
8a. Denied
8b. Approved
9. Researcher and relevant Executive Committee reps Sign Data Usage Agreement
10. Send data
11. Do research
12. Review prior to publication or presentation
12a. Destroy data
13. Research published, or presentation given.
14. Disseminate research

b) Data Requests

Participating Data Contributors (PDCs) have access to the Level 3 dataset after the initial CACDS "data cleaning" process on an annual basis. Level 3 data consist of de-identified, individual-level child data set from multiple PDCs. A formal approval request process is not necessary for internal reporting, but PDCs will be able to access the data by making a request ticket through the gacacds.com website. Requestors should go through the "Ask CACDS" portal, select the "data" button, then select "request new dataset." The submitted ticket will be reviewed by staff from the CACDS Administering Agency and once approved the requestor will be notified.

If the data will be used for reports to be distributed externally, with the Governor's Office, or the Office of Planning and Budget, the PDC will inform the CAA and the relevant PDC agencies, any agency whose data has been used in the reports, at least 15 days prior to distribution. The PDC will do this via email and include details such as major findings, methodology, and data used in the report. A copy of the report will be provided to the CAA and relevant PDC agency or agencies.

Third Party Managed Users planning on using Level 3 data on behalf of a PDC, through a contract agreement with the agency, will have to submit additional data agreement forms and sign the Data Usage Agreement. The CAA will also inform other agencies about the project and research questions.

Section 7:CACDS Technology and Data

a) CACDS Technology Overview

The core of Georgia's CACDS is a data warehouse — a relational database residing on Microsoft SQL Server 2016. The web portal was developed in ASP.NET using C#, Visual Studio 2015. The extract, transforming, load (ETL) tasks for importing and transforming the various source data are accomplished through Microsoft SSIS packages. Standard reporting is built on the Microsoft SSRS model. Business analytic tools and ad hoc analysis are accomplished using Microsoft Power BI along with a series of multidimensional Excel Cubes connecting to the aggregated statistical data.

i. Unique Identifiers

Through the GA Awards de-duplication process, primary functions of the system are to connect multiple data sources for an individual child and to assign a unique identifier (UID). GA Awards utilizes selected child demographic information (e.g., First Name, Last Name, Date of Birth, SSN, etc.) for an algorithm to create the UID and for matching purposes. However, none of the demographic information is used in the UID number. These UIDs are never duplicated. The UID is vital to produce an unduplicated child count for participating early childhood programs.

ii. CACDS Data Integration and Alignment

Early childhood data are received from multiple external sources. The child-level information is matched with other program data and a master child record is created. Program specific information is merged into an enrollment table along with provider and physical site data in which the service is provided. Aggregation removes the uniqueness and summarizes the program and demographic data for analytical and reporting purposes.

iii. User Portal

The CACDS website (www.gacacds.com) is the public-facing component of the application and allows all users to access supporting documents, view standard reports, customize the standard reports, communicate with a help desk, and submit data requests.

iv. Identified Researcher Portal

Georgia's CACDS also includes a limited-access portal that is utilized only by the CACDS Administering Agency (CAA), CAA Manager, and the Identified Researcher (IR) Portal. It allows IRs to access multiple data levels from PDCs. After completing a training, IRs will be able to access data sets and custom reports using a unique User ID that will be assigned to them.

v. Dashboards and Research Reports

In conjunction with the Executive Committee and Leadership Team, the CAA created public dashboards, which utilize CACDS data to display key indicators that track state or community progress in meeting key goals and benchmarks. The dashboards were completed in Fall 2018.

b) Data Use/Data Ownerships

The CAA and Executive Committee are committed to ensuring that CACDS policies and procedures comply with FERPA and the highest standards for data security. The following subsections discuss many of the standards.

i. Data Privacy and Confidentiality

The PDCs retain ownership of the data from their programs and data systems. In this context, "ownership" is understood to refer to the role assumed by a state agency in the creation, maintenance, and custody of the data pertaining to a program operated by that agency. CACDS acts as a broker providing agency data to those authorized to receive it under the terms of the DSAs and MOUs. PDCs authorize retrieval of agreed-upon data elements only. CACDS does not retrieve data elements not specified in the MOUs.

CACDS complies with all applicable National Institute of Standards and Technology (NIST) Security Controls required for state and federal compliance. These controls are listed in standard SM 17-001.

ii. Data Breaches

If a data breach is suspected or occurs within the Georgia CACDS application, the CAA will notify the Executive Committee and the impacted PDCs immediately, investigate the incident, and develop a plan to prevent future threats or incidents. The CAA/CACDS support team will present a report on its findings to the Executive Committee within three days of the incident that includes recommendations for corrective action, if necessary.

If a data requester, institution, or PDC experiences a data breach involving CACDS data, the party must report it to the CACDS Executive Committee within three business days from the date the party becomes aware of the breach or unauthorized data disclosure. The CAA/CACDS support team, in conjunction with the Executive Committee, may require the party involved in the breach to do any or all the following:

- Investigate and respond to the CACDS Executive Committee regarding any potential or actual breach and unauthorized data disclosure.
- Promptly resolve any problems identified by the investigation, if necessary.
- Submit a corrective action plan with steps to prevent any future unauthorized disclosures.
- Require that all the data files be destroyed immediately.

Failure to comply with these steps will result in restricted access to CACDS data in the future.

iii. Small Cell Suppression

The Georgia CACDS web portal builds upon and contains guidelines from the National Center for Education Statistics that address appropriate protection of children's identities in presenting statistical data.

As a general rule, no cell, table, or text discussion in any report or other publications provided by CACDS or by a data requester can include information that pertains to less than five children. Based on the specifics of the analysis, additional disclosure avoidance techniques, such as aggregation across categories, may be requested.