

# **\*\*Table of Contents\*\***

## **Forms**

1. Application for Federal Assistance (SF-424)	e1
2. Standard Budget Sheet (ED 524)	e5
3. SF 424B - Assurances Non-Construction Programs	e7
4. Disclosure of Lobbying Activities	e9
5. ED 80-0013 Certification	e10
6. Dept of Education Supplemental Information for SF-424	e11

## **Narratives**

1. Project Narrative - (Project Narrative - Project Abstract...)	e12
Abstract	e13
2. Project Narrative - (Project Narrative - Project Narrative...)	e14
Narrative_LDS	e15
3. Project Narrative - (Project Narrative - Appendix A, Optional Attach.....)	e43
Appendix A	e44
4. Project Narrative - (Project Narrative - Appendix B Resumes of Key P.....)	e47
Appendix B	e48
5. Project Narrative - (Project Narrative - Appendix C Current Status o.....)	e51
Appendix C	e52
6. Project Narrative - (Project Narrative - Appendix D Letters of Support...)	e56
Letters of Support	e57
7. Budget Narrative - (Budget Narrative - Budget Justification...)	e61
Budget Narrative	e62
8. Budget Narrative - (Budget Narrative - ED 524 Section C Spreadsheet...)	e71
Budget Spreadsheet	e72

This application was generated using the PDF functionality. The PDF functionality automatically numbers the pages in this application. Some pages/sections of this application may contain 2 sets of page numbers, one set created by the applicant and the other set created by e-Application's PDF functionality. Page numbers created by the e-Application PDF functionality will be preceded by the letter e (for example, e1, e2, e3, etc.).

**Application for Federal Assistance SF-424**

Version 02

* 1. Type of Submission	* 2. Type of Application:* If Revision, select appropriate letter(s):	
<input type="checkbox"/> Preapplication	<input checked="" type="checkbox"/> New	
<input checked="" type="checkbox"/> Application	<input type="checkbox"/> Continuation	* Other (Specify)
<input type="checkbox"/> Changed/Corrected Application	<input type="checkbox"/> Revision	

* 3. Date Received:	4. Applicant Identifier:
12/4/2009	

5a. Federal Entity Identifier:	* 5b. Federal Award Identifier:
	N/A

**State Use Only:**

6. Date Received by State:	7. State Application Identifier:

**8. APPLICANT INFORMATION:**

\* a. Legal Name: Georgia Department of Education

* b. Employer/Taxpayer Identification Number (EIN/TIN):	* c. Organizational DUNS:
586002042	806743159

**d. Address:**

* Street1:	205 Jesse Hill Drive
Street2:	
* City:	Atlanta
County:	Fulton
State:	GA
Province:	
* Country:	USA
* Zip / Postal Code:	30334

**e. Organizational Unit:**

Department Name:	Division Name:
Office of Technology Services	Information Technology

**f. Name and contact information of person to be contacted on matters involving this application:**

Prefix:	* First Name:	Robert
Middle Name:		

\* Last Name: Swiggum

Suffix:

Title: Georgia LDS Project

Organizational Affiliation:

Georgia Department of Education

\* Telephone Number:

(404)869-1011

Fax Number:

(404)651-9503

\* Email: RSWIGGUM@DOE.K12.GA.US

**Application for Federal Assistance SF-424**

Version 02

**9. Type of Applicant 1: Select Applicant Type:**

A: State Government

Type of Applicant 2: Select Applicant Type:

Type of Applicant 3: Select Applicant Type:

\* Other (specify):

**10. Name of Federal Agency:**

U.S. Department of Education

**11. Catalog of Federal Domestic Assistance Number:**

84.384A

CFDA Title:

Statewide Longitudinal Data System Recovery Act Grants

**\* 12. Funding Opportunity Number:**

NA

Title:

N/A

**13. Competition Identification Number:**

Title:

**14. Areas Affected by Project (Cities, Counties, States, etc.):**

Entire State of Georgia

**\* 15. Descriptive Title of Applicant's Project:**

Georgia LDS Project

Attach supporting documents as specified in agency instructions.

**Attachment:**

Title :

File :

**Attachment:**

Title :

File :

**Attachment:**

Title :

File :

**Application for Federal Assistance SF-424**

Version 02

**16. Congressional Districts Of:**

\* a. Applicant: GA

\* b. Program/Project: All

Attach an additional list of Program/Project Congressional Districts if needed.

**Attachment:**

Title :

File :

**17. Proposed Project:**

\* a. Start Date: 4/1/2010

\* b. End Date: 12/31/2012

**18. Estimated Funding (\$):**

a. Federal	\$ 14515480
b. Applicant	\$
c. State	\$
d. Local	\$
e. Other	\$
f. Program	\$
Income	
g. TOTAL	\$ 14515480

**\* 19. Is Application Subject to Review By State Under Executive Order 12372 Process?**

a. This application was made available to the State under the Executive Order 12372 Process for review on .

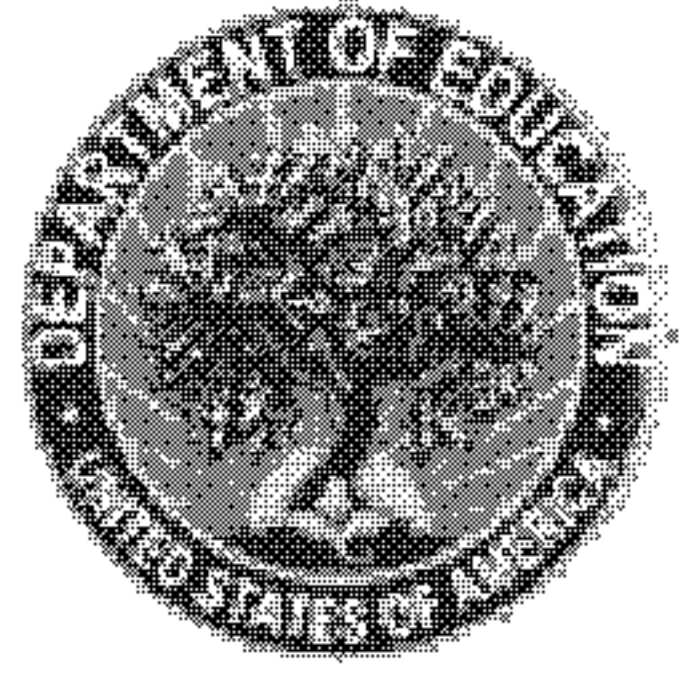
b. Program is subject to E.O. 12372 but has not been selected by the State for review.

c. Program is not covered by E.O. 12372.

**\* 20. Is the Applicant Delinquent On Any Federal Debt? (If "Yes", provide explanation.)**

Yes  No





**U.S. DEPARTMENT OF EDUCATION**  
**BUDGET INFORMATION**  
**NON-CONSTRUCTION PROGRAMS**

OMB Control Number: 1894-0008

Expiration Date: 02/28/2011

Name of Institution/Organization:  
Georgia Department of Education

Applicants requesting funding for only one year should complete the column under "Project Year 1." Applicants requesting funding for multi-year grants should complete all applicable columns. Please read all instructions before completing form.

**SECTION A - BUDGET SUMMARY**  
**U.S. DEPARTMENT OF EDUCATION FUNDS**

Budget Categories	Project Year 1(a)	Project Year 2 (b)	Project Year 3 (c)	Project Year 4 (d)	Project Year 5 (e)	Total (f)
1. Personnel	\$ 2,500,162	\$ 2,500,162	\$ 2,143,012	\$ 0	\$ 0	\$ 7,143,336
2. Fringe Benefits	\$ 999,998	\$ 999,998	\$ 857,148	\$ 0	\$ 0	\$ 2,857,144
3. Travel	\$ 5,000	\$ 5,000	\$ 5,000	\$ 0	\$ 0	\$ 15,000
4. Equipment	\$ 1,500,000	\$ 0	\$ 0	\$ 0	\$ 0	\$ 1,500,000
5. Supplies	\$ 600,000	\$ 400,000	\$ 200,000	\$ 0	\$ 0	\$ 1,200,000
6. Contractual	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
7. Construction	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
8. Other	\$ 600,000	\$ 600,000	\$ 600,000	\$ 0	\$ 0	\$ 1,800,000
9. Total Direct Costs (lines 1-8)	\$ 6,205,160	\$ 4,505,160	\$ 3,805,160	\$ 0	\$ 0	\$ 14,515,480
10. Indirect Costs*	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
11. Training Stipends	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
12. Total Costs (lines 9-11)	\$ 6,205,160	\$ 4,505,160	\$ 3,805,160	\$ 0	\$ 0	\$ 14,515,480

**\*Indirect Cost Information (To Be Completed by Your Business Office):**

If you are requesting reimbursement for indirect costs on line 10, please answer the following questions:

(1) Do you have an Indirect Cost Rate Agreement approved by the Federal government?  Yes  No

(2) If yes, please provide the following information:

Period Covered by the Indirect Cost Rate Agreement: From: \_\_\_/\_\_\_/\_\_\_ To: \_\_\_/\_\_\_/\_\_\_ (mm/dd/yyyy)

Approving Federal agency:  ED  Other (please specify): \_\_\_\_\_ The Indirect Cost Rate is \_\_\_\_\_%

(3) For Restricted Rate Programs (check one) -- Are you using a restricted indirect cost rate that:

Is included in your approved Indirect Cost Rate Agreement? or,  Complies with 34 CFR 76.564(c)(2)? The Restricted Indirect Cost Rate is \_\_\_\_\_%



**U.S. DEPARTMENT OF EDUCATION**  
**BUDGET INFORMATION**  
**NON-CONSTRUCTION PROGRAMS**

OMB Control Number: 1894-0008

Expiration Date: 02/28/2011

Name of Institution/Organization:  
 Georgia Department of Education

Applicants requesting funding for only one year should complete the column under "Project Year 1." Applicants requesting funding for multi-year grants should complete all applicable columns. Please read all instructions before completing form.

**SECTION B - BUDGET SUMMARY**  
**NON-FEDERAL FUNDS**

Budget Categories	Project Year 1(a)	Project Year 2 (b)	Project Year 3 (c)	Project Year 4 (d)	Project Year 5 (e)	Total (f)
1. Personnel	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
2. Fringe Benefits	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
3. Travel	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
4. Equipment	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
5. Supplies	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
6. Contractual	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
7. Construction	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
8. Other	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
9. Total Direct Costs (lines 1-8)	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
10. Indirect Costs	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
11. Training Stipends	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0
12. Total Costs (lines 9-11)	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0	\$ 0

## ASSURANCES - NON-CONSTRUCTION PROGRAMS

Standard Form 424B (Rev.7-97)

Public reporting burden for this collection of information is estimated to average 15 minutes per response, including time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding the burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to the Office of Management and Budget, Paperwork Reduction Project (0348-0040), Washington DC 20503.

PLEASE DO NOT RETURN YOUR COMPLETED FORM TO THE OFFICE OF MANAGEMENT AND BUDGET. SEND IT TO THE ADDRESS PROVIDED BY THE SPONSORING AGENCY.

**NOTE:** Certain of these assurances may not be applicable to your project or program. If you have questions, please contact the awarding agency. Further, certain Federal awarding agencies may require applicants to certify to additional assurances. If such is the case, you will be notified.

As the duly authorized representative of the applicant, I certify that the applicant:

1. Has the legal authority to apply for Federal assistance, and the institutional, managerial and financial capability (including funds sufficient to pay the non-Federal share of project cost) to ensure proper planning, management, and completion of the project described in this application.
2. Will give the awarding agency, the Comptroller General of the United States, and if appropriate, the State, through any authorized representative, access to and the right to examine all records, books, papers, or documents related to the award; and will establish a proper accounting system in accordance with generally accepted accounting standards or agency directives.
3. Will establish safeguards to prohibit employees from using their positions for a purpose that constitutes or presents the appearance of personal or organizational conflict of interest, or personal gain.
4. Will initiate and complete the work within the applicable time frame after receipt of approval of the awarding agency.
5. Will comply with the Intergovernmental Personnel Act of 1970 (42 U.S.C. "4728-4763) relating to prescribed standards for merit systems for programs funded under one of the 19 statutes or regulations specified in Appendix A of OPM's Standards for a Merit System of Personnel Administration (5 C.F.R. 900, Subpart F).
6. Will comply with all Federal statutes relating to nondiscrimination. These include but are not limited to: (a) Title VI of the Civil Rights Act of 1964 (P.L. 88-352) which prohibits discrimination on the basis of race, color or national origin; (b) Title IX of the Education Amendments of 1972, as amended (20 U.S.C. "1681-1683, and 1685-1686), which prohibits discrimination on the basis of sex; (c) Section 504 of the Rehabilitation Act of 1973, as amended (29 U.S.C. '794), which prohibits discrimination on the basis of handicaps; (d) the Age Discrimination Act
9. Will comply, as applicable, with the provisions of the Davis-Bacon Act (40 U.S.C. "276a to 276a-7), the Copeland Act (40 U.S.C. '276c and 18 U.S.C. "874) and the Contract Work Hours and Safety Standards Act (40 U.S.C. " 327-333), regarding labor standards for federally assisted construction sub-agreements.
10. Will comply, if applicable, with flood insurance purchase requirements of Section 102(a) of the Flood Disaster Protection Act of 1973 (P.L. 93-234) which requires recipients in a special flood hazard area to participate in the program and to purchase flood insurance if the total cost of insurable construction and acquisition is \$10,000 or more.
11. Will comply with environmental standards which may be prescribed pursuant to the following: (a) institution of environmental quality control measures under the National Environmental Policy Act of 1969 (P.L. 91-190) and Executive Order (EO) 11514; (b) notification of violating facilities pursuant to EO 11738; (c) protection of wetlands pursuant to EO 11990; (d) evaluation of flood hazards in floodplains in accordance with EO 11988; (e) assurance of project consistency with the approved State management program developed under the Coastal Zone Management Act of 1972 (16 U.S.C. "1451 et seq.); (f) conformity of Federal actions to State (Clear Air) Implementation Plans under Section 176(c) of the Clear Air Act of 1955, as amended (42 U.S.C. "7401 et seq.); (g) protection of underground sources of drinking water under the Safe Drinking Water Act of 1974, as amended, (P.L. 93-523); and (h) protection of endangered species under the Endangered Species Act of 1973, as amended, (P.L. 93-205).
12. Will comply with the Wild and Scenic Rivers Act of 1968 (16 U.S.C. "1721 et seq.) related to protecting components or potential components of the national wild and scenic rivers system.
13. Will assist the awarding agency in assuring compliance

of 1975, as amended (42 U.S.C. " 6101-6107), which prohibits discrimination on the basis of age; (e) the Drug Abuse Office and Treatment Act of 1972 (P.L. 92-255), as amended, relating to nondiscrimination on the basis of drug abuse; (f) the Comprehensive Alcohol Abuse and Alcoholism Prevention, Treatment and Rehabilitation Act of 1970 (P.L. 91-616), as amended, relating to nondiscrimination on the basis of alcohol abuse or alcoholism; (g) " 523 and 527 of the Public Health Service Act of 1912 (42 U.S.C. " 290 dd-3 and 290 ee 3), as amended, relating to confidentiality of alcohol and drug abuse patient records; (h) Title VIII of the Civil Rights Act of 1968 (42 U.S.C. ' 3601 et seq.), as amended, relating to nondiscrimination in the sale, rental or financing of housing; (i) any other nondiscrimination provisions in the specific statute(s) under which application for Federal assistance is being made; and (j) the requirements of any other nondiscrimination statute(s) which may apply to the application.

7. Will comply, or has already complied, with the requirements of Titles II and III of the uniform Relocation Assistance and Real Property Acquisition Policies Act of 1970 (P.L. 91-646) which provide for fair and equitable treatment of persons displaced or whose property is acquired as a result of Federal or federally assisted programs. These requirements apply to all interests in real property acquired for project purposes regardless of Federal participation in purchases.
8. Will comply, as applicable, with the provisions of the Hatch Act (5 U.S.C. "1501-1508 and 7324-7328) which limit the political activities of employees whose principal employment activities are funded in whole or in part with Federal funds.

with Section 106 of the National Historic Preservation Act of 1966, as amended (16 U.S.C. '470), EO 11593 (identification and protection of historic properties), and the Archaeological and Historic Preservation Act of 1974 (16 U.S.C. "469a-1 et seq.).

14. Will comply with P.L. 93-348 regarding the protection of human subjects involved in research, development, and related activities supported by this award of assistance.
15. Will comply with the Laboratory Animal Welfare Act of 1966 (P.L. 89-544, as amended, 7 U.S.C. "2131 et seq.) pertaining to the care, handling, and treatment of warm blooded animals held for research, teaching, or other activities supported by this award of assistance.
16. Will comply with the Lead-Based Paint Poisoning Prevention Act (42 U.S.C. "4801 et seq.) which prohibits the use of lead- based paint in construction or rehabilitation of residence structures.
17. Will cause to be performed the required financial and compliance audits in accordance with the Single Audit Act Amendments of 1996 and OMB Circular No. A-133, "Audits of States, Local Governments, and Non-Profit Organizations."
18. Will comply with all applicable requirements of all other Federal laws, executive orders, regulations and policies governing this program.

**Signature of Authorized Certifying Representative:**

**Name of Authorized Certifying Representative:** Bob Swiggum

**Title:** Chief Information Officer

**Date Submitted:** 12/01/2009

### Disclosure of Lobbying Activities

Complete this form to disclose lobbying activities pursuant to 31 U.S.C. 1352

<b>1. Type of Federal Action:</b> <input type="checkbox"/> Contract <input checked="" type="checkbox"/> Grant <input type="checkbox"/> Cooperative Agreement <input type="checkbox"/> Loan <input type="checkbox"/> Loan Guarantee <input type="checkbox"/> Loan Insurance	<b>2. Status of Federal Action:</b> <input type="checkbox"/> Bid/Offer/Application <input checked="" type="checkbox"/> Initial Award <input type="checkbox"/> Post-Award	<b>3. Report Type:</b> <input type="checkbox"/> Initial Filing <input type="checkbox"/> Material Change <b>For Material Change only:</b> Year: 0 Quarter: 0 Date of Last Report:
<b>4. Name and Address of Reporting Entity:</b> <input checked="" type="checkbox"/> Prime <input type="checkbox"/> Subawardee Tier, if known: 0 Name: N/A Address: City: State: Zip Code + 4: -  <b>Congressional District, if known:</b>	<b>5. If Reporting Entity in No. 4 is a Subawardee, Enter Name and Address of Prime:</b> Name: Address: City: State: Zip Code + 4: -  <b>Congressional District, if known:</b>	
<b>6. Federal Department/Agency:</b> U. S. Department of Education	<b>7. Federal Program Name/Description:</b> Statewide Longitudinal Data Systems  CFDA Number, if applicable: 84.384A	
<b>8. Federal Action Number, if known:</b>	<b>9. Award Amount, if known:</b> \$0	
<b>10. a. Name of Lobbying Registrant</b> (if individual, last name, first name, MI): N/A Address: City: State: Zip Code + 4: -	<b>b. Individuals Performing Services</b> (including address if different from No. 10a) (last name, first name, MI): N/A Address: City: State: Zip Code + 4: -	
<b>11.</b> Information requested through this form is authorized by title 31 U.S.C. section 1352. This disclosure of lobbying activities is a material representation of fact upon which reliance was placed by the tier above when this transaction was made or entered into. This disclosure is required pursuant to 31 U.S.C. 1352. This information will be reported to the Congress semi-annually and will be available for public inspection. Any person who fails to file the required disclosure shall be subject to a civil penalty of not less than \$10,000 and not more than \$100,000 for each such failure.	Name: Robert Swiggum Title: Chief Information Officer Applicant: Georgia Department of Education Date: 11/30/2009	
<b>Federal Use Only:</b>		Authorized for Local Reproduction Standard Form LLL (Rev. 7-97)

## CERTIFICATION REGARDING LOBBYING

Certification for Contracts, Grants, Loans, and Cooperative Agreements.

The undersigned certifies, to the best of his or her knowledge and belief, that:

(1) No Federal appropriated funds have been paid or will be paid, by or on behalf of the undersigned, to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with the awarding of any Federal contract, the making of any Federal grant, the making of any Federal Loan, the entering into of any cooperative agreement, and the extension, continuation, renewal, amendment, or modification of any Federal contract, grant, loan or cooperative agreement.

(2) If any funds other than Federal appropriated funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this Federal contract, grant, loan or cooperative agreement, the undersigned shall complete and submit Standard Form - LLL, "Disclosure of Lobbying Activities," in accordance with its instructions.

(3) The undersigned shall require that the language of this certification be included in the award documents for all subawards at all tiers (including subcontracts, subgrants and contracts under grants, loans, and cooperative agreements) and that all subrecipients shall certify and disclose accordingly. This certification is a material representation of fact upon which reliance was placed when this transaction was made or entered into. Submission of this certification is a prerequisite for making or entering into this transaction imposed by section 1352, title 31, U.S. Code. Any person who fails to file the required certification shall be subject to a civil penalty of not less than \$10,000 and not more than \$100,000 for each such failure.

Statement for Loan Guarantees and Loan Insurance.

The undersigned states, to the best of his or her knowledge and belief, that:

If any funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee or any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this commitment providing for the United States to insure or guarantee a loan, the undersigned shall complete and submit Standard Form-LLL, "Disclosure of Lobbying Activities," in accordance with its instructions. Submission of this statement is a prerequisite for making or entering into this transaction imposed by section 1352, title 31, U.S. Code. Any person who fails to file the required statement shall be subject to a civil penalty of not less than \$10,000 and not more than \$100,000 for each such failure.

### APPLICANT'S ORGANIZATION

Georgia Department of Education

### PRINTED NAME AND TITLE OF AUTHORIZED REPRESENTATIVE

Prefix:                      First Name: Robert    Middle Name:

Last Name: Swiggum    Suffix:

Title: Chief Information Officer

Signature: \_\_\_\_\_    Date: 12/01/2009



# Project Narrative

## Project Narrative - Project Abstract

Attachment 1:

Title: **Abstract** Pages: **1** Uploaded File: **\\gtcrsan2\Falcon\Docs\am294\Desktop\IES\Abstract\_LDS.pdf**

## **Project Abstract: Georgia Longitudinal Data System (GLDS)**

The Georgia Alliance of Education Agency Heads (AEAH) was created by Governor Perdue in 2006 and consists of the state's seven education-related agencies. The Alliance agencies are: the Department of Early Care and Learning (Pre-K), Department of Education (K-12), Technical College System of Georgia, University System of Georgia, Georgia Student Finance Commission, Professional Standards Commission, and the Governor's Office of Student Achievement. The Alliance is charged with collaborating on policies and programs that will prepare our students for the opportunities and challenges of the 21st century. The AEAH has identified multiple goals and strategies for improving the quality of education in Georgia, focusing on preparation for postsecondary education and transition into the workforce. However, to measure the effectiveness of the Alliance objectives, longitudinal data are needed that currently do not exist in Georgia.

In March 2009, the Georgia Department of Education (GADOE) was awarded \$8.9 million from the Institute of Education Sciences Statewide Longitudinal Data Systems grant (Georgia's "Chronicle" project). These funds are primarily designed to automate, streamline, enhance, and manage the data collection and reporting activities for the state's K-12 system. Additionally, it provided funds to develop and prepare for the linking of data with other state education partners. This application will complement and build upon the work started under the previous IES grant. Additionally, this proposal aligns with the opportunities and expectations of the Race to the Top Grant for which Georgia will apply, and also with the outcomes and expectations of the College and Career Ready Policy Institute (CCRPI) of which Georgia is a participant

Presently we are performing ad-hoc data matching of a longitudinal-nature using manual extracts and rudimentary matching criteria. These matches are typically performed to meet singular reporting needs, are time consuming, and provide limited information from which to make informed decisions. A statewide longitudinal data system (LDS) is needed to effectively measure outcomes, as well as provide continuous improvement, transparency, and accountability. This system will also serve as a resource for the seven educational agencies to track a student's academic history, thus contributing toward a seamless educational system.

The outcomes of this project include (1) the design, development, and implementation of a robust student-level, dimensional, P-20 LDS linking student, teacher, and teacher-preparedness information; (2) advanced data matching algorithms to ensure accurate matching of students from multiple source systems; and (3) development of a Decision Support System that will provide primary user-level access to the library of reports and information needed for decision making.

# Project Narrative

## Project Narrative - Project Narrative

Attachment 1:

Title: **Narrative\_LDS** Pages: **28** Uploaded File: \\gtrcsan2\Falcon\Docs\am294\Desktop\IES\Narrative\_LDS.pdf

## **Project Narrative**

### **(a) Need for Project**

Georgia's application for grant support of the development and completion of a statewide longitudinal P-20 data system is the product of the combined efforts of the state's education agencies: Georgia Department of Education (GADOE); University System of Georgia (USG); Technical College System of Georgia (TCSG); Department of Early Care and Learning (DECAL); Georgia Student Finance Commission (GSFC); Professional Standards Commission (PSC); and the Governor's Office of Student Achievement (GOSA). These agencies have formed an Alliance of Educational Agency Heads (AEAH), established in 2005, that shares five announced and agreed upon common educational goals: (1) Increase high school graduation rate, decrease high school drop-out rate, and increase postsecondary enrollment rate; (2) Strengthen teacher quality, recruitment and retention; (3) Improve workforce readiness skills; (4) Develop strong educational leaders, particularly at the building level, and; (5) Improve the SAT/ACT scores of Georgia students. The need for this support can be described in terms of supplementing Georgia's manifold efforts to improve its current education data collection efforts.

Georgia's efforts to improve its educational system throughout the entire educational P-20 to work pipeline depend on accurately and systematically describing in data both the educational process and the results of that process: student progression through the system and the ultimate success of students in endeavors for which they were prepared by the educational process. Data to inform those descriptions come from information collections primarily designed to support the day-to-day educational enterprise. These short-run business uses of data are important, but the sum of daily activities also leads to the ultimate end of education – student preparation. What follows is a description of Georgia's existing data collection efforts, our strengths and current data improvement activities, and the weaknesses of the data that requires additional work that must be done for us to develop, construct, and later maintain a longitudinal data system that supports local, state and national needs. Finally, Georgia is a participant in the College and Career Ready Policy Institute (CCRPI) which is a competitively awarded grant designed to provide technical assistance to the nation's leading states to help them move from good to great in the following policy areas: goals, assessment, accountability, supports and interventions, pathways and models, and data. Within the CCRPI, Georgia has chosen to focus primarily on goals, assessment, accountability, and data; with the guiding principle that data is central to all other areas. The goals of the CCRPI align very well with the goals of the work proposed with this IES grant application as well as the Race To The Top program. The three opportunities are interrelated and complement one another regarding P20 longitudinal data systems. We continue with a brief summary of what our longitudinal data

system will be able to do if our request for funds is approved under the terms of this application.

### *An Overview of Georgia's Existing Data Collections*

The GADOE has worked in recent years to improve its K-12 data collections as a means of improving student outcomes. The most pressing data needs have been associated with the requirements on state student achievement mandated by the No Child Left Behind Act (NCLB), especially year-to-year progress on test results and measures related to high school graduation rates and school attendance. These K-12 data collections also support the state's full-time-equivalent-student based Quality Basic Education Act funding formula and provide data for program identification with a subsequent basis for the evaluation of state programs. Collections come from student information systems used in the 182 school systems and their 2,000+ schools. The student and business information systems used in Georgia's public schools are not standardized, but instead are purchased from a variety of vendors whose implementation in particular schools is frequently customized for the school.

The improvement of these K-12 collections will be supported by the development of Chronicle, the GADOE project designed to increase the frequency of data collections to better ensure the quality and timely availability of K-12 data. The main goal of Chronicle will be the production of formative data to improve instruction and the application of educational services. In performing these functions, data will also be generated to better calculate dropout rates and in tracking K-12 students, including students who may leave the state. Chronicle will also generate higher quality data that will become part of the P-20 data within the State Longitudinal Data System (SLDS).

The USG collects extensive institutional data on student enrollment, course-taking, remediation, and program participation as well as broader programmatic and faculty information. Its data collections are supported by its institutional student information system and various business systems. Standardization of software (Banner) exists in most of the USG except in the flagship institution, the University of Georgia, which uses its own legacy system. The USG has historically been able to provide system wide high school feedback summary reports to Georgia high schools on the progress of high school graduates within the USG.

The TCSG similarly collects extensive institutional data on student enrollment, course-taking, remediation and program participation as well as programmatic and faculty information. TCSG data collections are supported by common applications like Banner in the case of student information. While both TCSG and USG rely on Banner software, different data definitions and coding require data normalization in order to obtain comparable data. TCSG has also attempted to assess the linkage between participation in TCSG programs and the labor market.

The DECAL administers Georgia's voluntary pre-kindergarten for four-year-olds through a variety of public and private providers. Data collections within DECAL will provide data related to the nature and quality of services delivered to students in the program as well as student and teacher information. DECAL has chosen to adopt the GADOE's Georgia Testing Identifier (GTID) to facilitate student matches and ensure availability of data to assess the effects of its programs and program providers, including the qualifications of pre-kindergarten teachers and the various curricula offered and their relationships to subsequent student progress. This decision will aid in matching these students with students in the public education system.

The GSFC collects complete transcript data for all Georgia public and private high school graduates. This collection of data includes course-taking and grade histories for all high school grades (grades 9-12), and includes school grading and grade-weighting practices. The agency, through its GAcademy411.org mentor site, sends transcript data on behalf of students to Georgia colleges to which the students have applied. As required by state law, GSFC calculates grade point averages for all students using a common calculation method for purposes of determining student eligibility for the HOPE Scholarship, Georgia's merit-based scholarship. The GSFC also collects information on the use of state scholarship and grant programs by students within the state in both public and private institutions and for a few out-of-state colleges with which the state has a reciprocity agreement.

The PSC, as the agency responsible for teacher licensure, collects information on the educational background of all teachers licensed by Georgia, as well as information on all teacher preparation programs within the state. Additionally, PSC (with GADOE) collects information on Georgia public school teacher assignments, training and experience and salaries through the CPI (Certified/Classified Personnel Information) collections.

The GOSA is the agency responsible for collecting data from agencies and publishing data reflecting student achievement in the state of Georgia's P-20 educational system for accountability purposes. Its contributions to data collection include a linkage to data contained in the National Student Clearinghouse (NSC), which includes student attendance data for approximately 92% of all college students in the United States. By linking graduates to NSC data, GOSA has been able for the first time to publish systematic data on postsecondary participation by college for Georgia's public school graduates.

Given our current collections, the shared vision for a Georgia LDS involves the collection of data from business systems within education agencies (often obtained from school or school system level business systems) and the cleansing and storage of that data in agency-specific data warehouses. Data from the data warehouses will periodically be transferred to a P-20 data staging operation where data will be transformed and normalized for inclusion in the P-20 data warehouse. The state's longitudinal data system may be described as the entire chain of

collections from school to P-20 data warehouse, including links to labor and employment data. Reports and data extracts to provide information for education policy decision-makers or to summarize the educational process will come from the standardized data in the P-20 warehouse.

*Shortcomings in Current Data Collection and Reporting Efforts that must be Addressed in Building a Longitudinal Data System*

There is no agreed upon P-20 data model for the Georgia LDS, and the development of such a model will require an initial audit of potential data elements to be used in that system and a gap analysis of what remains to be done to improve data.

The elements and data element definitions may be well-known to Georgia agency data collectors and to those who supply the elements at the school/system/data provider level, but their presence, consistency over time and accuracy must be thoroughly explored before elements can be used as the basis of a longitudinal data system. The result of this data audit process will be to construct a P-20 model that is consistent with internal needs and with state and national reporting requirements. The results may also establish the need for modifications in the current data collections. This gap analysis will form the basis for a continuous data improvement model, including training for agency staff, by which data will be periodically audited in the light of changing requirements, such as data to support RTTT strategies. This process will be necessary to meet the requirements of capabilities two<sup>1</sup> and six<sup>2</sup> for state longitudinal data systems.

*There is not a standard agreed upon method of matching students who move from or transition to another educational level or from education providers in one sector to another.*

Past efforts to link data from various educational agencies have involved idiosyncratic data matching methods that have yielded mixed results. Within the GADOE, the standardization of identifiers for student testing purposes (the GTID, an algorithm based identifier created by the GADOE) has made significant headway in tying together data for Georgia students who stay in

---

<sup>1</sup> “The system must facilitate and enable the exchange of data among agencies and institutions within the State and between States so that data may be used to inform policy and practice. Such a system would support interoperability by using standard data structures, data formats, and data definitions to ensure linkage and connectivity among the various levels and types of data.”

<sup>2</sup> “The system must ensure the quality and integrity of data contained in the system”

the K-12 public sector, but there is still a residual problem in that the GTID is not used by students in all sectors and at all levels. The most common identifier, the student Social Security Number, is still used by most of the agencies involved in data collection efforts, including the GADOE, although other identifiers are used as well. Because of this, and because the LDS will require student records that join together data generated in different agencies, places and times, research on the most effective and efficient way to link student data must be undertaken so that a series of matching techniques may be used to build an LDS that identifies and links students with a very low percentage of unlinked data. This student data matching methodology can include data elements such as school, school system, address, date of birth, course-taking matches, and other data elements currently collected at the agency level. Once the matching process is specified, managers of the LDS can begin to put in place a series of data cleansing routines that will be necessary to periodically update and maintain the system, and a means of identifying students with a unique identifier for use in identifying matched students for the purpose of consolidating student records and for no other purpose. This single statewide student identifier is a required element<sup>3</sup> in state longitudinal data systems as described in the America COMPETES Act. Similarly, the ability to match students across the divides in the educational pipeline into postsecondary education is another required element that is currently incomplete in Georgia that will be satisfied upon completion of this work<sup>4</sup>.

Finally, a uniform method of tracking students from the educational system into employment and labor markets must be developed in order to assess the effects of education on the workplace, on the affluence of students related to their educational preparation and experiences, and on their needs for subsequent re-training and continuing education. Work in these areas that has previously been done in Georgia (in TCSG and by USG) has not been systematic enough for a true longitudinal data system with better student matching methods.

---

<sup>3</sup> “ A unique statewide student identifier that does not permit a student to be individually identified by users of the system (except as allowed by Federal and State law).”

<sup>4</sup> “ Data that provide information regarding the extent to which students transition successfully from secondary school to postsecondary education, including whether students enroll in remedial coursework.”

*There is need for research on the best and most secure way to provide reports and data from the P-20 data warehouse to the various users: parents, teachers, school and system administrators, policy-makers, state and federal agencies, researchers and the general public.*

Once the P-20 data model is developed and implemented and student matches have allowed for the development of consolidated student records a method for data reporting and data transfer must be developed. These methods should reflect the data needs of the various stakeholders in the educational system while safeguarding the privacy of students and teachers whose data is in the system. These methods will be in addition to any data transfer methods for formative purposes contained in the agency warehouses which are part of the LDS, like the data used by GADOE's Chronicle Student Profile system. The perfection of the student data matching methodology will enable postsecondary results data to be made available to inform both state and local education decision-making, a capability required of state longitudinal data systems<sup>5</sup>.

*Summary: The Functioning of Georgia's Longitudinal Data System Upon Final Construction Under the Terms of this Grant Application*

The state of Georgia has been tracking student achievement and success through interagency agreements for some years. Investigating and assessing how policies and policy changes affect post-secondary enrollment rests on our ability to track progression down to the student level. In addition to student level trend analysis, the state has developed a teacher preparation database to support the ability to analyze in-state post secondary teacher education programs.

This grant allows the infrastructure of the longitudinal data system (LDS) to be established as a series of operation data store (ODS) extractions into a dimensional data base. Currently the various education agencies utilize secure flat-file transfer utilities (FTU) and varied load processes for matching students between agencies from pre-K to post-secondary. Even with the long willingness to share data among the agencies, there has been no seamless way to merge and match data. Measuring outcomes through the education systems and into the labor force via the transferred flat file process requires vast resources. The matching process needed to uniquely identify students once their information is transferred between and among the various agencies is convoluted and tedious. This process begs the establishment and

---

<sup>5</sup> "The system must enable data to be easily generated for continuous improvement and decision-making, including timely reporting to parents, teachers, and school leaders on the achievement of their students."

implementation of a unique identifier to streamline and improve the process and facilitate more timely integration and interfacing of information among these agencies. The time needed to match students will be reduced with an identifier unique to Georgia. Such a unique identifier should adhere to any federal, state or locally mandated rules or legislation for protecting the student's identify.

The LDS will require data management to ensure quality and validity of the data. Incorporating data from seven agencies into one system for student tracking and reporting will require the creation a cleansing area for managing and ensuring data quality as well as a reporting layer to facilitate end user reporting. To begin, the warehouse the data elements need to be inventoried and normalized. Upon normalizing the data, an extraction process with a matching algorithm can be applied to place the data in the proper dimensions. Between departments a standardization of elements, either through a transfer or reference table, and appropriate data and metadata management software would need to be acquired and utilized to allow for a truly functioning longitudinal system.

Creating a LDS will allow for longitudinal studies to be performed. The analyses now performed with the departments sharing data consists more of cross sectional and repeated measures, but is not a true longitudinal analysis. These analyses would measure student transitions between grades and departments, teacher effectiveness, and provide data for decision making. Reporting from one system will provide the State the ability to identify and follow federal metrics in a timely manner and adjust policy to improve performance. Parents, teachers, and administrators will be granted differing roles through a business intelligence tool to follow the achievement of their students.

The increased complexity to comply with federal, state, and local districts' reporting needs and analysis due to separate systems has pushed the need for a unified data warehouse in Georgia. In addition to reporting, automated exchanges of data are needed for timeliness and cleansing. Collecting the data at certain time periods allows for scheduled reports to run for validation purposes. The current manual method of matching data between departments limits the ability to validate data and fix issues rapidly. Validation currently occurs only when running reports, thus some reports are withheld past deadlines due to not knowing the data issues until seeing the results.

*Some of the key features that will be supported by Georgia's proposed Longitudinal Data System that are not supported by existing resources and/or activities include the following.*

The creation of systematic data on course-taking patterns for graduates of public and private high schools and their effects on postsecondary progress variables such as hours attempted, hours earned, grades, choice of major, progression within the postsecondary system, etc.

Effects of teachers related variables (certification level, experience, preparation, qualities related to the preparation program, etc.) on student achievement within the P-12 and on postsecondary progress variables such as hours attempted, hours earned, grades, choice of major, progression within the postsecondary system, etc.

Linkages between educational experiences (participation in programs, coursework, etc.) and eventual workforce-related variables.

Please See Appendix C for more information on Georgia's current status regarding the capabilities and elements outlined in Section IV of the RFA.

## **(b) Project Outcomes Related to System Requirements and Implementation**

The specific outcomes envisioned for the Georgia Longitudinal Data System project are:

Outcome 1: Development of the Data System

1.1: Data Audit and Analysis

1.2: Data Model and Data Architecture

1.3: Extraction-Transformation-Load (ETL) process

Outcome 2: Improvement of Data Matching Algorithm Across Agencies

Outcome 3: Create a Decision Support System for the GLDS

3.1 Creation of initial high-level public reports that address the high school to college transition.

3.2 Document user reporting needs and functionality for all types and levels of user.

3.3 Evaluate business intelligence, dashboard, and reporting tools.

3.4 Build reporting layer access and security.

3.5 Build new reports according to user needs documented in Outcome 3.2.

### ***Outcome 1: Development of the Data System***

#### ***1.1 Data Audit and Analysis***

The GLDS will be fed from elements originally collected from business users of data that are transmitted at a polled point-in-time collection. While the points-in-time vary according to the collection type, their frequency determines the effort necessary to clean and square the data with previously collected data. For example, K-12 data collections depend on data obtained

from school districts which themselves depend on school level data entry and recording of data elements (in the public school example) or directly from the schools themselves (in the private school example). Data that conflicts with previously collected data must be reconciled.

Since agencies do not typically oversee data entry, the transfer of data from the business users through the district to the state agency typically involves data quality control checks called “edits” which generate error and/or warning messages designed to either reject data elements that do not conform to the proper data or content standards imposed by the data-receiving agency or to prompt the business users to confirm that the element which generates the warning message is accurate. These edits are designed to ensure data quality, but their continued consistent use rests on the importance of the edit to later actual uses of the data. Thus, edits that allow final data collection and subsequent data acceptance to proceed while generating an “error” or “warning” message for the attention of the business user/data supplier may result in data satisfactory for the immediate business use but ultimately unsatisfactory for use in a longitudinal data system. A related problem is the “correction” of data in response to an “error” or “warning” for the purposes of the collection without modifying data in the core business system. This may lead to a repeat of the initial error/warning message in subsequent collections. Within the K-12 data collections within the Georgia Department of Education, these tasks will be performed as part of the Chronicle work effort already supported by an IES grant to support GADOE’s longitudinal data system development.

The interconnections among data collected at different points-in-time must be reconciled in longitudinal data systems, and the discovery and documentation of those methods of data reconciliation must be completed as part of the construction of such a system. To discover these methods, each agency participating in the longitudinal data system must perform a detailed data audit of each of its data collections, including every element collected. Each agency audit must contain a systematic categorization and listing of values received for each element together with and in addition to the analysis of permissible values and data collection standards that comprise typical data dictionaries. These audits should be applied to each collection that could potentially supply data to the longitudinal data system, and so should extend to past collections over a period of years including the most recent collections. Problems with missing data or incorrectly specified elements in each named collection must be noted as a part of this process. After the initial agency level data audits are complete, each agency must determine whether existing processes for resolving data conflicts will suffice for providing data for a longitudinal data system or whether additional steps will be necessary to improve data quality in prospective agency data collections.

Special attention must be paid in the analysis of data obtained by agencies in collections from business users to the creation of derived or computed data elements. These are a variety of

composite elements that are composed of lower level elements. Examples of these are rates, full-time-equivalent student counts, grade-point-averages, and the like. If collected elements are themselves composite elements, then the decision rules underlying the specification of the coded element must be made explicit in the audit documentation.

The agency data audit process will result in extensive documentation that assesses the results of existing data collections. The sum of these agency data audits will constitute the baseline analysis for creating the elements, collected, derived or calculated, which will populate the longitudinal data system warehouse. The process for creating these LDS elements will be discovered during the analysis of what transformational data processes are necessary to complete the population of the LDS warehouse. The resulting documentation will establish the exact formulae that generate each of these elements, in the case of each derived or calculated element, and the exact coding logic required for each non-calculated element in the system. The transformational steps, when complete, will ensure a robust and consistent set of longitudinal data elements.

Finally, the audit and analysis of existing data must be compared to data necessary to measure student achievement (Race to the Top, etc.) growth over time to ensure that requirements for educational reform measurement data are met.

In summary, the outcomes of this data audit and analysis process are:

- Detailed agency data audits describing the elements collected over time, their content, use, and descriptions of any issues related to their quality, especially those that call for modifications in data collections;
- Detailed agency descriptions of derived or calculated elements, including the formulae used incorporating elemental data obtained from collections;
- Analysis of what collected, derived or calculated measures may be included in the longitudinal data system and for what years;
- Discovery and creation of a process of data transformation and subsequent loading based on initial data populating the LDS plus data from continuing collections to be housed in the LDS
- Comparison of historical longitudinal data to future reform data needs with recommendations for identifying improvements to existing collection content or frequencies to meet those needs.

## *1.2: Data Model and Data Architecture*

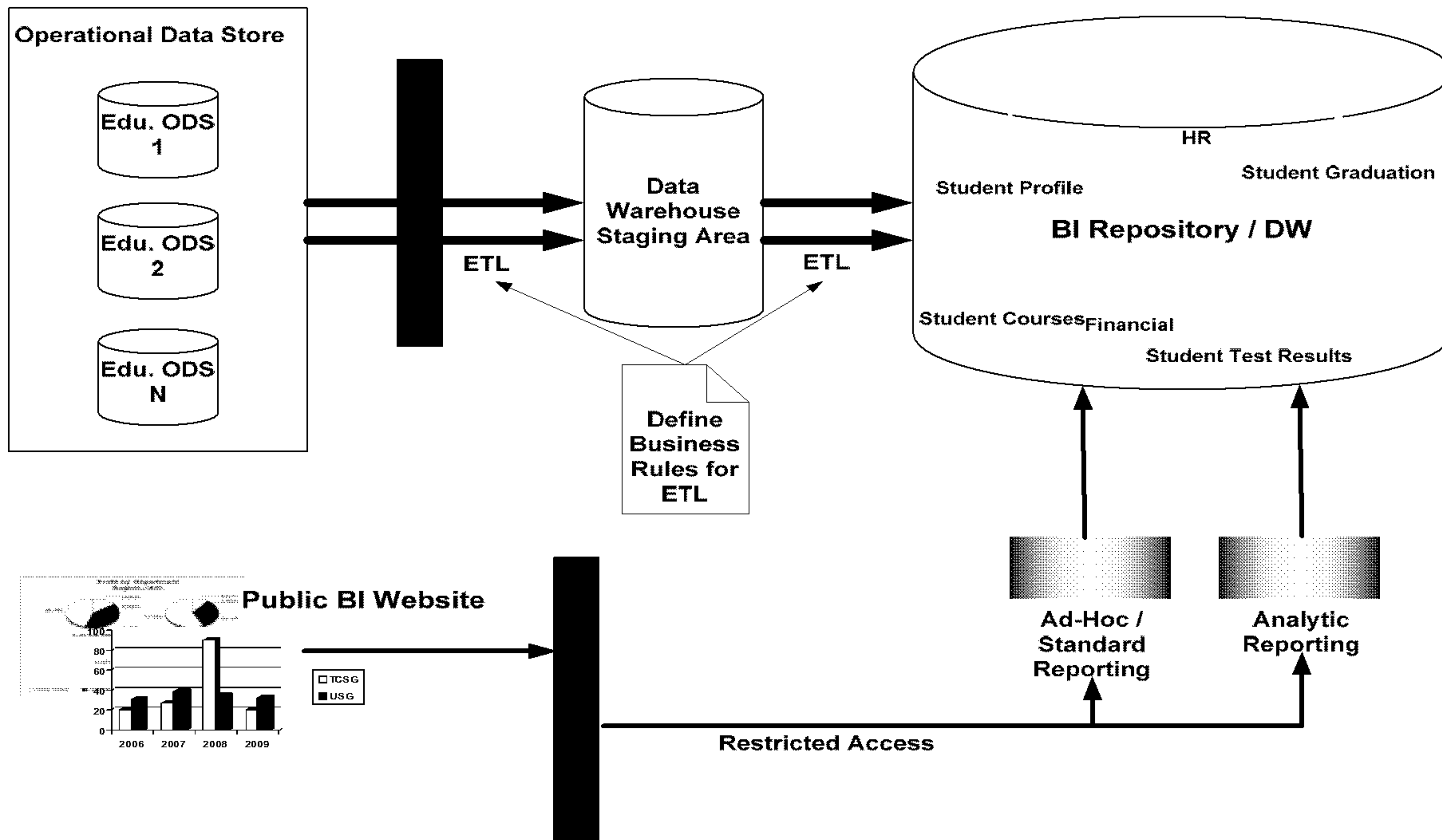
Data warehousing is a viable solution for an organization to provide meaningful access to the strategic information. Building a data warehouse and the associated business intelligence applications requires integrating data from multiple data sources to create a platform for running business analytical applications. Data Warehousing facilitates subject oriented longitudinal data analysis while meeting performance expectations.

Any data warehouse is only as useful as the design. The architecture should first understand the end users' requirements, business rules, and the design it for reporting and analysis needs. Warehouses are repositories of data for business intelligence tools and reporting tools sit upon. Different design approaches can be taken, normalized or dimensional, and the methodologies can also differ: top down or bottom up. Before the LDS can be complete a survey of all the reporting agencies needs to occur to establish the exact purposes, intentions, and questions to be answered.

A data warehouse allows integration of data into a series of subject-area data marts designed around specific business processes and decision needs. The design of these subject areas will be driven by the business process and reporting needs from various stakeholders. In order to design an extensible data model for the warehouse subject areas, it is very crucial to learn and document the business processes within various operational data stores, as well as documenting specific reporting needs. A clear set of business rules must be developed for each of the business process identified. Data update frequencies must be established. Obviously, the data marts must meet stakeholder needs for easily accessible, readily available, relevant data and information. More specifically, the data marts are intended to:

- Utilize data to better facilitate decision making by institutional and System leadership;
- Assist institutions and the System to be more accountable to various communities;
- Provide both pre-defined reporting and the ability for ad hoc querying;
- Ability to report across data marts (subject areas);
- Implement a self-service user interface to all data marts for all consumers – general public or users with role-based access;
- Provide executive dashboards.

Data warehouse implementations generally involve data extraction, transformation, and loading (ETL) processes that are defined for each of the business process. Data is cleansed, extracted, transformed and loaded into data warehouse staging area. Data within data warehouse staging area is then integrated with another set of ETL routines into meaningful star schemas.



A reporting repository is created either on low level star schemas or second level aggregated star schemas, based on the reporting/aggregation/presentation needs. The reporting repository facilitates slicing-dicing/drilling-aggregating data over various dimensions and facts, allowing various sorts and presentation parameters. Business intelligence reporting tools access either the reporting repository or the database star schemas to allow generation of executive dashboards, canned standard reports, ad-hoc /self-service reporting. OLAP cubes built on top of the star schemas allow trend analysis and what-if analysis. These analyses are very useful for generating metrics for various measures.

Two major approaches exist in data warehousing, Kimball and Inmon. Kimball's design (bottom up) attached data marts via a bus structure. Connecting all the marts by common elements allowed for ease of querying the marts in unison. The marts could be located on one server or multiple but the connection allowed for a virtual warehouse because the data would be pulled from the marts. The data marts are directly loaded from the ODS tables depending on the requirements.

Inmon's approach (top down) seeks a transfer of data into on centralized location. The data marts in this approach are created from the warehouse and do not make the warehouse; created marts exist for reporting purposes of the differing business areas. The data are loaded with ODS during the extraction process.

Again, until the customer's are surveyed and reporting requirements understood both approaches are possibilities. Understanding the customers' needs has to occur prior to

deciding on a design. The architecture chosen will allow the customer to utilize business intelligence tools, reporting devices, OLAP cubes, and data mining. The design will be a combination of styles based on the business requirements.

### *1.3: Extraction-Transformation-Load (ETL) process*

The process of collecting data from the different agencies will follow common warehousing technique. The first step is the extraction process. The integration of all of the disparate systems across is the first hurdle in creating a warehouse. Utilizing complex scripts heterogeneous data is extracted (copied) and placed into staging tables. Each sources system will have distinct characteristics requiring finesse and resources to manage and integrate the data into the tables. This process includes the creating of a data map, which will be critical in determining the validity of the data pulled. Documentation during this period is essential to inform the developers of the location of the data. The location and values of the elements allow the examination of the source data to reduce redundancy, integration, and track the point-of-origin.

Implementation of an ETL and data store processes where the data transfer, editing and cleansing process does not unnecessarily involve agencies or their transactional collection systems is required for smoothness, transparency, and efficiency. If the ETL does not meet these requirements it will burden the agencies and allow uncleansed data to be populated. One method of assessment for the ETL is the use of metadata.

The transform step uses a set of logical rules or functions with the extracted data before loading into the final location. Often during this step translation tables (Xref tables) are utilized to normalize data from the varying systems. This step changes the data to a useable or measurable unit. Data quality check points exist in this step to ensure consistent, complete, usable data are loaded. This can occur through an edit process using established business rules. Documentation of these rules are critical for establishing metrics of adherence and defining the data elements.

A successful data warehouse design is attributable to well-defined business requirements that are closely aligned with the information and decision support needs of the business user. Once these are defined and created a load process moves the data from the staging area into the warehouse. The architecture and end business needs shape how the load process will work, whether the warehouse employs dimension or fact tables.

Metadata is defined as “data about data”. A comprehensive approach to metadata management allows an organization to standardize, map and store its information about data in terms of data sources, content, business rules and access authorizations in a centralized repository, thus documenting the enterprise’s information resources across non-integrated

systems. It does so preferably in easily understood terminology. The end result is that the organization can integrate, share, centrally manage, and leverage information across its transactional database silos. Administration of metadata typically also allows for documenting evolutionary changes in existing data definitions over time as well as creation of new data definitions as they emerge in the organization.

Data definitions and parameters must be established systematically. Data must be consistent for the LDS to be useful. Data definitions should consider standards already in place at the national and state levels. Data models should be developed that are flexible enough to allow meaningful comparisons and to establish and maintain integrity regardless of the source. A needs assessment by the agencies should be performed to address what data are needed but not collected or what is unnecessary to collect. Streamlining the data collected and refining the needs would save resources.

## **Outcome 2: Improvement of Data Matching Algorithm Across Agencies**

### *Creating Student Identifiers for the Georgia Longitudinal Data System*

Maintaining student data over time and across institutional divides requires a method for tracking students. This tracking requires that students be identified at a variety of times and places, even when different methods for identifying students are used by different agencies or at different times. Absent the use of a perfectly reproducible unique identifier based on unalterable biometric elements, other primary and secondary identifiers must be used. Existing primary student identifiers for data contained in the various agencies which will contribute data to Georgia's LDS include: Georgia Testing Identifier (GTID); Social Security Numbers; School- or System-Generated IDs; and Name (F-M-L). Other secondary data elements that can be used in ascertaining or verifying data matches include, but may not be limited to: school; school system; address, city, county, zip code and state; parent name, address, city, county zip code and state; telephone number; date of birth; gender; race/ethnicity; course(s) taken; grade level or grade year; link to teacher.

While student matching has been performed for a variety of purposes in Georgia, the matching techniques used have been idiosyncratic and have yielded mixed results. To create and maintain a longitudinal data system that tracks students, a standardized technique must be developed that yields the highest possible matching algorithm so that student data can follow the student across time and institutional divides. The goal of this outcome is to analyze, from the many possible matching techniques, the best stepwise progression of activities to obtain the best match possible. This method or methods will be employed any time new data processed for addition to the longitudinal data system

In order to attack this problem, we will attempt to set a high standard for our data matches: we propose, as a goal, to match all data collected at any point in an agency collection to a master data file for a student. Before creating a master data file for a newly-encountered student, we will attempt to verify the unique status of each “new” student by creating ongoing patterns of algorithm-based matching with existing master data files that show no apparent match in newly acquired data using standard student identifiers. The unique identifier (the “Master Student Record,” or MSR) for each student for use in the LDS will be linked, within the secure system, with all other identifiers used for that student, along with links to the data obtained for that student in all of the data collections in the student’s history. When students are matched through methods other than the standard student identifiers used by business users and data collections, a history of how each student was linked to each collection will be established so that the secondary method can be used for the student in the future if the primary matching method fails.

To monitor our progress towards this goal, we will set a number of matching statistics by which the effectiveness and efficiency of our matching methods may be summarized. These process indicators will give users of the data system an appropriate assurance as to the quality of the data generated by the system. The indicators must show the match of students across collections within agencies as well as showing matching statistics across institutions and educational divides.

Of particular interest in matching students is the transition from one type of institution to another, most importantly the transition from high school to postsecondary education. The longitudinal data system will create identifiers for “hard cohorts”, so that students who have received a high school graduation credential (this is what makes the cohort “hard”, in the sense that their status is absolutely defined at one point in time) are noted with unique identifiers that describe their point of graduation with regard to both time of graduation and place (high school). While these cohorts do not describe all high school students, the remaining non-graduates will be matched with GED completers and with program participants in the Technical College System of Georgia, with private college Tuition Equalization Grant data, and with the National Student Clearinghouse data to see if they can be matched. Unmatched student master files not in the “hard cohort” will also be matched against employment and other workforce data. Data in this hard cohort, which will include both public and private high school graduates, will also be matched against previous K-12 unmatched data files, thus lowering the frequency of current unmatched K-12 data records.

All master student data files will also be matched annually with employment data from the Georgia Department of Labor. Identification of elements from that matched employment data will also be included as part of the longitudinal data system.

In summary, the products of this student identifier creation process are:

- Establishing data matching statistics which characterize the effectiveness and efficiency of our historical matching techniques;
- Using primary and secondary identification elements to devise standardized data matching techniques for matching students in the GLDS;
- Devising a common master student data file for each student encountered in any single educational data collection so that all data collected on a student can be linked back to the student;
- Creation of master data files for all students beginning with the historical data identified in the data audit and analysis outcome as the first year for inclusion in the longitudinal data system;
- Thorough documentation of all data match findings, techniques and specification for the addition of additional data in periodic collections for the GLDS.

### **Outcome 3: Create a Decision Support System for the GLDS**

#### *3.1 Initial Dashboards and Reports*

The success of this initiative depends on an early demonstration of what the GLDS will provide when completed. This is necessary to engage key stakeholders and garner the short- and long-term support to secure the funding and political will necessary to sustain the GLDS beyond the initial grant-funded startup period. To that end, an initial series of high level public-access reports will be developed and made available by November, 2010. These dashboard and indicator reports will focus on key transition points in the education pipeline. These reports will represent an “early win” for the GLDS, and will help build support among various stakeholder groups for sustainability.

These initial public reports will also play a critical role in understanding and documenting future user needs. These reports will provide the critical stimulus to initiate and promote user engagement in the process of future report development. This is discussed further in Outcome 3.2.

For the initial demonstration phase of the GLDS, high school and post-secondary student data will be linked to produce integrated high school to college transition reports. The student transition reports will include, for example: high school courses; high school graduation and diploma types; HOPE scholarship eligibility; post-secondary enrollment (TCSG institutions or USG institutions); remediation in post-secondary; grades in introductory college mathematics and English courses; and college retention (persistence). The linked data will be loaded into

simple data marts. Expanded reporting features, dashboard, indicators, scorecards, and information presentation will be developed and made available to the public via a web portal in November, 2010.

Georgia has already done preliminary work in this area. For years, education agencies have manually shared data for purposes of reporting on student transitions. These reports have been targeted to very specific, circumscribed report requirements, have been of limited use, and have required considerable manual matching and processing. Very little useful, readily accessible information has been made available to policy makers and the public.

To develop these initial reports and dashboards, existing data from DOE, GSFC, TCSG, GOSA and USG will need to be matched and linked. The manual matching process used in previous data sharing has varied between agencies. The matching methodology has not been adequately documented. The accuracy and extent of the matching has also not been documented and assessed. The matching required for these initial reports and dashboards will serve as the initial documentation and analyses necessary for the development of the Student Record Matching Algorithm described above in Outcome 2.

Additionally, this early reporting demonstration will serve as the initial analysis of reporting needs and functionality. To ensure development of high quality reports that provide critical information on student transitions, a sound understanding of the informational needs is critical. Careful documentation and analysis of reporting needs will be conducted to provide the detail necessary to develop the reports, dashboards, and scorecards.

While the reports that will be produced in this initial phase of the GLDS may sound impressive, it does not constitute the formation of an SLDS. The data matching process would still be a manual process. Although the reports will have easy access, have some extended functionality, and will convey some critical information, they would nevertheless still be essentially static reports limited to specific subsets of data for specific years. Ongoing maintenance and usefulness of this method of reporting is very limited, inefficient, ineffective, and not acceptable considering what is available once an SLDS is implemented.

### *3.2 Document user reporting needs and functionality for all types and levels of users.*

For any SLDS to be truly useful, it must meet the information needs of all of its user types and levels. A deep understanding of what users need and how they need to access it can only be derived through direct involvement of all levels and types of users. To this end, extensive research of user needs will be undertaken. Focus groups, interviews, and other means of documenting user needs will be undertaken. This phase of the project will gather information

from all identified GLDS user groups. Specifically, we will seek input from each group on questions such as: what information do they need to access; what level of data manipulation do they need; what type of user interface do they need; what type of report functionality (e.g., drill down / up, disaggregation, etc.) do they need; what output functionality do they need; what level of granularity do they need (e.g., some users will only need access to aggregated data while analysts / power users may need access to anonymized unit level data); and so on.

The initial dashboard and reports developed in Outcome 3.1 will serve a critical role in this process. Frequently, users have ideas of what they want but they need something to react to. Having examples of the types of reports that they would be have access to will spur much more extensive and detailed discussions of how a decision support system should be developed to meet their current and anticipated needs.

Working with user groups to understand reporting needs will begin in October, 2010, and will continue through January, 2011.

In addition to the direct input from constituent users, a detailed review and analysis of existing cross-sector reporting requirements will be conducted. This will include examinations of federal, state, and local reporting requirements and their relation to existing data. This analysis should be completed by December, 2010.

Based on the data elements and reporting requirements analyses, a comprehensive picture of data needs will be developed. Documentation and compilation of reporting requirements analyses should be completed by February, 2011. These analyses of reporting requirements will be passed to the full-scale report development objective in Outcome 3.5.

### *3.3 Evaluate business intelligence, dashboard, and reporting tools and web-based presentation tools.*

Each agency has implemented its own solution to analyzing data and generating reports and information. These include various combinations of business intelligence tools, in-house-developed tools, and off-the-shelf report development packages. Implementing the GLDS will require investing in decision support system / business intelligence tools and web-based presentation tools. Based on the reporting requirements, data model, and data architecture developed in the aforementioned objectives, the State will develop an RFP for a business intelligence solution. The Alliance will appoint a committee to assist in the RFP development and to evaluate the resulting proposals. The RFP will be completed and posted by March, 2011. Evaluation of proposals and products will take place between April, 2011, and May, 2011. Recommendations will be made to the Alliance in June, 2011 for final selection and approval.

### *3.4 Build reporting layer access and security.*

All reporting functionality will be accessible only at the reporting layer (not directly from the data warehouse). The policies specifying who may have access to the GLDS will shape the general procedures for granting access. Reporting access will be role-based at multiple levels ranging from public access to restricted “power-user” data analyst access. Based on the user needs, the appropriate reporting layer data structures will be developed (e.g., cubes, data marts, etc.) to provide access only to specified data at specified levels of aggregation. These structures will be populated with data from the warehouse and procedures for refreshing the reporting structures will be developed and implemented.

Based on the specific web-based reporting / presentation tools and the specific dashboard / business intelligence / analysis tools selected in Outcome 3.3, additional features will likely be required in the reporting layer.

Security will be a critical component. A streamlined user account creation / maintenance system will be developed. Based on the individual’s account, access to specific reports and reporting structures will be granted. A secure web interface will be developed and implemented as the vehicle to accessing the reports and information. An access and usage monitoring / auditing process will also be developed and implemented. Additional security features will be implemented as needed based on an extensive security audit of the system as it is developed.

The building of security procedures and the reporting layer should begin by May, 2011, and should be completed by September, 2011.

### *3.5 Build new reports according to user needs documented in Outcome 3.2.*

Al the reporting layer is developed, full scale report development will begin. Based on the reporting requirements documented in Outcome 3.2, a broad catalog of reports will be developed to provide access to information for all constituent user groups. These reports will include the full breadth of the data contained in the GLDS and will go far beyond the initial dashboard reports, including workforce data from the Georgia Department of Labor.

Additionally, the reports will include data obtained from the National Student Clearinghouse (NSC). In 2009, Georgia entered into a contract with the Bill and Melinda Gates Foundation to work on the Gates NSC Pilot Project, which will improve the match rate between Georgia’s high school graduate records and NSC’s database, to analyze the postsecondary enrollment patterns

of Georgia's high school graduates and to produce accessible reports on this topic. This work is a natural fit for inclusion in the GLDS project.

The primary goal of this report catalog development is to provide extensive reports on students at transition points: transition from pre-k into K-12; transition through K-12 into post-secondary; transition from K-12 directly into the workforce; transition from post-secondary into the workforce. Detailed analysis and reporting capability will be made available to examine factors associate with successful transitions throughout the education to workforce pipeline. Full-scale feedback reports will be developed and made available.

A critical component of report development will be the establishment of policies and procedures for report verification and validation. Additionally, necessary and sufficient documentation (metadata and user documentation) must be developed.

Report development will begin by August, 2011. The initial report catalog will be completed by December, 2012. However, report development is an ongoing process and will continue well beyond the grant period.

### **(c) Timeline for Project Outcomes**

#### **Outcome 1: Development of the Data System**

##### *Documentation and Analysis of Data Elements and Reporting Requirements*

The first step in building the GLDS is the design and development of a sound, scalable data model. Understanding what data are available, adequate, and sufficient is a critical prerequisite. Likewise, an understanding of the informational and reporting requirements for the GLDS must be analyzed and understood up front in order to develop an appropriate data model and to develop the GLDS architecture.

Data elements analysis – Each state agency contributing data has its own set of data elements and data definitions. A comprehensive review of all data elements will be undertaken to: identify which elements are needed to meet the informational and reporting needs; determine the quality of each element; determine the granularity of each element; consistency of data element definitions, formats, and coding; and determine which years of data should be considered for inclusion. The compilation and analysis of all data elements should be completed by September, 2010.

Reporting requirements analysis – To develop an understanding of the types of reporting needs, user types / levels, necessary data, and other information needs, a detailed review and

analysis of existing cross-sector reporting requirements will be conducted. This will include examinations of federal, state, and local reporting requirements and their relation to existing data. Additionally, reporting and information needs will be solicited from participating agencies, state-level policymakers, and other diverse audiences to further expand the reporting requirements analysis. Documentation and compilation of reporting requirements analyses should be completed by September, 2010. The compilation and analyses of reporting requirements also constitutes the ongoing process of expanding the reporting capability of the GLDS.

Synthesis – Based on the data elements and reporting requirements analyses, a comprehensive picture of data needs will be developed. This synthesis will identify the data needed, which elements are currently available for inclusions in the GLDS, any critical data elements not already collected that should be collected, and any derived measures necessary for inclusion. This synthesis should be completed by September, 2010.

#### *Develop Data Model and Data Architecture*

Based on the data elements and reporting requirements analyses and synthesis, the data model will begin development November 2010. The data model will incorporate data from all seven Georgia education agencies. Additionally, the model will include Georgia Department of Labor data. This stage of development is absolutely critical and will have an intense focus. The model should accommodate the existing data and reporting needs but must also be designed to be scalable to include additional data and to meet additional reporting requirements as needed. The initial data model should be completed by January, 2011. The data model will be validated and vetted by all participating agencies by March, 2011.

#### Timeline:

Data Audit and Analysis: April-September 2010.

Data Model and Data Architecture: November 2010-March 2011.

Extraction Transformation Load (ETL) process:

Extraction: December 2010-March 2011.

Transformation: March-July 2011.

Load: May-September 2011.

### **Outcome 2: Improvement of Data Matching Algorithm Across Agencies**

The successful implementation of a matching algorithm will require significant research and development work during the first phase of the GLDS project. A team of business analysts, researchers, programmers, technical writers, and agency personnel will review best practices used by other states or systems, identify possible routines that would work for Georgia, write the programming required for the routines, test and validate the results of the various options, and ultimately propose an advanced matching algorithm that will be implemented as the project continues. It will be very important for this team to coordinate with the partnering agencies to fully understand the data elements that are available, conditional use of those data, and special requirements that may exist. This work will happen concurrently as the first phase of the GLDS is being developed. Implementation of the SRMA is anticipated by January 2011.

Timeline:

Improvement of Data Matching Algorithm Across Agencies: July 2010-January 2011.

### **Outcome 3: Create a Decision Support System for the GLDS**

Timeline:

3.1 Creation of initial high-level public reports that address the high school to college transition: April-November 2010.

3.2 Document user reporting needs and functionality for all types and levels of user: October 2010-February 2011.

3.3 Evaluate business intelligence, dashboard, and reporting tools: February 2011-June 2011.

3.4 Build reporting layer access and security: May 2011-September 2011.

3.5 Build new reports according to user needs documented in Outcome 3.2: August 2011-December 2012.

Please see Appendix A.1, "Georgia Longitudinal Data System Project Timeline" for more detail on the timeline for these subtasks. Please see Appendix C for more information on how activities supported by this grant will be coordinated with activities supported by other grants.

### **(d) Project Management and Governance Plan**

#### *Project Location*

The Governor's Office of Student Achievement, as the state's independent education accountability agency will house the LDS Director, Project Manager and staff. Each agency will maintain responsibility for their data collections and will submit data defined by the processes

described in this section to the GLDS. Additional staff will be identified at each agency to act as liaisons to the LDS staff on behalf of their respective agencies.

The project's governance structure will consist of three levels. At the highest level, the Data Governance Board will consist of the agencies' Chiefs of Staff, and will have the following responsibilities:

- Set/enforce policies for data collection and access
- Give final approval for outside data requests
- Resolve data issues/conflicts
- Promote use of the LDS data
- Decide other governance processes

The Data Governance Board will develop bylaws and will have final authority over interagency management of the state LDS.

Reporting to the Data Governance Board will be the Data Management Committee, which will consist of the agency CIOs. The responsibilities of the Data Management Committee will include:

- Act as the stewards of the data
- Approve new data elements to be collected
- Improve data quality
- Improve data accuracy
- Improve understanding of data
- Improve data use
- Present data requests to Data Governance Board for final approval

The Data Management Committee will make recommendations for final approval to the Data Governance Board. Two subcommittees will report to the Data Management Committee: the Information Technology Group and the Research Group. The Information Technology Group will consist of technical experts from the agencies, and their main responsibility will be to review requests to resolve issues or to answer requests from non-agency researchers that concern technical details. The Research Group will function as the Institutional Review Board

for the GLDS. This group will be a rotating committee consisting of agency researchers and outside researchers (university faculty, for example). Their main responsibilities will be to review outside requests for methodological issues. This group may also encourage and guide research from within the agencies and from outside researchers.

Georgia's seven separate education agencies came to agreement on this structure in fall 2009. Please see Appendix A.2, titled "Governance Structure," and Appendix A.3, titled "Governance Principles," for more detail on this topic.

### *Management Controls*

The Georgia Department of Education currently follows the best practices established by the Project Management Institute (PMI) in the development of the Chronicle project, and will continue that practice with this project. Those practices include:

- Project Initiation: includes clearly defining the project's scope and expected outcomes;
- Project Planning: includes clearly defining the activities and resources that will deliver the final outcomes;
- Project Execution and Control: includes executing on the project plan, reporting on project status and ensuring project controls;
- Project Closeout: includes administrative closure and logistics, as well as user(s) accepting the final deliverables.

The Project Director will have authority for the project. The Project Director will work with the appropriate staff members from the State education agencies during the course of the project. Ultimately, the Project Director will sit on the Data Management Committee, as described above. The Project Manager will coordinate the day-to-day operations of the system from within the Governor's Office of Student Achievement.

### *Partnering Agencies*

The agencies involved in the creation of the GLDS are Georgia's seven state education agencies, led by the Alliance of Education Agency Heads (AEAH). The support and involvement by all agencies will be required to ensure the success of the project. The Alliance is comprised of leaders from the following agencies:

- State Superintendent of Schools of the Georgia Department of Education
- President of the Georgia Student Finance Commission

- Chancellor of the Board of Regents of the University System of Georgia
- Commissioner of the Department of Early Care and Learning
- Executive Director of the Governor's Office of Student Achievement
- Executive Secretary of the Georgia Professional Standards Commission
- Commissioner of the Technical College System of Georgia

*Data Sharing, Access, and Security Policies and Procedures*

Policies and procedures governing data sharing, access, and security must be developed at the beginning of the GLDS. The AEAH constitutes the State's data governance body. It has ultimate authority and approval powers covering the GLDS. The AEAH will appoint a "Data Governance Board" which will be comprised of representatives from all participating agencies (described above). This group will develop draft documents, policies, and procedures regarding the ongoing data governance body composition, data sharing, data access, report development, and security. These drafts will be submitted to the Alliance for approval.

The Data Governing Board will develop the following:

- Data sharing. Memoranda of understanding will be developed to specify the terms, conditions, uses, and security measures surrounding each agency providing data for inclusion in the GLDS. A Data Transfer Memorandum of Understanding (MOU) has been developed as part of Georgia's Race to the Top application.
- Data access. Access to the actual data warehouse will be strictly controlled. All reporting functionality will be accessible only at the reporting layer. Policies specifying who may have access to the data warehouse will be developed separate from policies determining reporting access. Reporting access will be role-based at multiple levels ranging from public access to restricted "power-user" data analyst access. Policies and procedures will be needed for laying out how users will be designated, terms and conditions they must abide by, what data they may access, and so on.
- Report requests and report development. Agencies and policy makers will have ongoing needs for additional reports to be developed and made available through the GLDS. The governance body will develop procedures for how such requests will be submitted, evaluated, and prioritized. Additionally, the governance body will develop procedures for validating and verifying the accuracy and completeness of reports prior to making them available through the GLDS.

- Access to data for research purposes. Researchers will make frequent requests for access to data and information for research purposes. The governance body must develop policies and procedures for how such requests will be submitted and how the requests will be evaluated and prioritized. Policies must be developed to ensure compliance requirements, confidentiality and FERPA requirements, and distribution and publication requirements.
- Security policies. The governance body will also develop policies and procedures ensuring the security of data and information contained in and derived from the GLDS . These policies would include physical security, electronic security, FERPA compliance, and other conditions determined by the governance body.

The governance body will complete its draft policies and procedures and submit them to the Alliance for approval in July, 2010. The Alliance will approve final policies and procedures in August, 2010. At that time, the full governance body will be convened and the policies and procedures will be officially implemented.

Please see Appendix D for letters of support from participating agencies and others.

### **(e) Project Staffing**

#### *Project Director*

The GADOE's Chief Information Officer will serve as the Project Director. In this role, the Project Director will receive regular reports from the project staff, and will provide oversight, including ensuring that the project is aligning appropriately with the Chronicle project and with Race to the Top. (0.10 FTE)

#### *Project Manager*

The Project Manager will be responsible for managing the agreement between the various State agencies involved in this project, contract negotiation and management, supervision of the project team, and providing updates on progress to the Project Director, as well as the Data Governing Board and Data Management Committee. The Project Manager will have points of contact within each of the other State education agencies and will oversee and coordinate day-to-day project activities. This person is currently not on the staff of any agency; GOSA plans to identify and hire the GLDS Project Manager as soon as possible but no later than March 2010. (1.0 FTE)

During the grant period, the Project Manager will directly manage 12 additional positions within GOSA to complete the project. The responsibilities of staff will shift according to the timeline included in Appendix A (Georgia Longitudinal Data System Project Timeline). Staff directly reporting to the GLDS Project Manager include:

- IT Support Specialist (1)
- Business analyst (3)
- Database programmer (3)
- Technical writer (1)
- Database Architect (1)
- Web Developer (1)
- Business Intelligence Developer (2)

Over the life of the project, the AEAH anticipates that permanent staff levels needed to initiate the GLDS project on state funding will be higher at the beginning of the project, and will be more than the levels needed to maintain the GLDS project by the end of the award.

In addition, each agency will identify one or more Agency Liaisons who will be dedicated specifically to this project and working with each of their agencies to support implementation of the GLDS project. Some of Georgia's education agencies have a greater need to expand and improve their existing systems in order to effectively contribute to the outcomes described in this application, and these staff levels reflect the varying levels of need:

University System of Georgia: 7

Georgia Student Finance Commission: 4

Technical College System of Georgia: 4

Professional Standards Commission: 3

Georgia Department of Education: 1

Department of Early Care and Learning: 1

Department of Labor: 1

*Infrastructure Requirements:*

Hardware and software infrastructure requirements will be completed by October, 2010. By October, 2010 the data storage and reporting requirements will be completed and an accurate assessment of required infrastructure can be determined.

*Hardware Requirements:*

Initial discussions are occurring among the AEAH about physical housing of the GLDS. The hardware infrastructure will require four environments: Development, QA, Pre-production and Production.

*Software Requirements:*

Anticipated software/licensing requirements are as follows:

- Database software
- Business Intelligence software
- ETL software
- Web development software

# Project Narrative

## Project Narrative - Appendix A, Optional Attachments

Attachment 1:

Title: **Appendix A** Pages: **3** Uploaded File: \\gtrcsan2\Falcon\Docs\am294\Desktop\IES\Appendix A\_LDS.pdf

## Appendix A: Optional Attachments

### Appendix A.1 Georgia Longitudinal Data System Project Timeline

#### *Project Initiation*

---

Hire GLDS Director	12/2009-3/2010
--------------------	----------------

#### *Outcome 1: Development of the Data System*

---

1.1: Data Audit and Analysis	4/2010-9/2010
1.2: Data Model and Data Architecture	11/2010-3/2011
1.3: Extraction-Transformation-Load (ETL) process	12/2010-9/2011
Extraction	12/2010-3/2011
Transfer	3/2011-7/2011
Load	5/2011-9/2011

#### *Outcome 2: Improvement of Data Matching Algorithm Across Agencies*

---

Improve Data Matching Algorithm	7/2010-1/2011
---------------------------------	---------------

#### *Outcome 3: Create a Decision Support System for the GLDS*

---

3.1 Creation of initial high-level public reports that address the high school to college transition.	4/2010-11-2010
3.2 Document user reporting needs and functionality for all types and levels of user.	10/2010-2/2011
3.3 Evaluate business intelligence, dashboard, and reporting tools.	2/2011-6/2011
3.4 Build reporting layer access and security.	5/2011-9/2011

