



Education Data Warehouse Serves Important Function; Project Planning and Management Need Strengthening

at a glance

Florida's Education Data Warehouse plays an important role in the state by enabling policymakers, educators, and researchers to track student progress from prekindergarten through graduate school. The warehouse was the first of its kind to be developed and is nationally recognized. However the warehouse has not fully implemented tools to improve stakeholders' access to education data. In addition, the Department of Education needs to better plan and manage its efforts to improve the data warehouse and source database systems to avoid duplication of effort and unnecessary costs, and to ensure that the state's information needs are met.

Scope

As directed by the Florida Legislature, this report examines the Department of Education's data warehouse and the department's information technology planning and management efforts for its major source database systems.

Background

Beginning in 2000, the Legislature appropriated funding to the Department of Education to create and maintain a statewide data warehouse to improve the state's ability to compile information about Florida's students and educational institutions. The department and the Board of Governors collect a wide variety of data from

Florida's school districts, community colleges, technical centers, and universities. These data include student demographic, course enrollment, attendance, assessment, financial aid, and program completion information; institutional finance and facilities information; and staff demographic and payroll information.

Reliable student and institutional data play an important role in improving Florida's education system. Analyses of these data can enable policymakers and educators to gauge the effects of education reform initiatives as well as meet federal reporting requirements. For example, both Florida's A++ Education Performance Accountability System and the federal No Child Left Behind Act require longitudinal analyses to assess student learning gains over time.

Florida is nationally recognized for its education data capabilities, including the state's Education Data Warehouse which serves as a model for other states developing such data warehouses. Florida was the first state to integrate its data across all public education sectors: K-12 schools, community colleges, and universities. The federal government recently cited Florida's ability to track a student's education from childhood through college while calling for higher state-level education standards and increased comprehensive education data. Florida is also one of six states to have all 10 elements of a statewide education data system as measured by the Data Quality Campaign, a national, collaborative effort to encourage and

support the improvement, availability, and use of high-quality education data.

Historically, Florida’s primary systems for collecting education data have been aligned with the state’s major educational sectors: K-12 public schools, technical centers, community colleges, and the State University System. Many of these systems were developed separately and implemented some 20 years ago under separate governing bodies using different technical platforms, software, data definitions, and reporting processes. While these systems provided useful information for each educational sector, it was difficult to track students longitudinally over time to evaluate the impact of various education reform efforts.

To address this problem, the Florida Legislature has appropriated over \$20.9 million to the department to develop, implement, and operate the Education Data Warehouse. As shown in Exhibit 1, the department has spent nearly \$7.3 million of these funds for contracted services to develop and implement the warehouse, with the remaining \$13.6 million used to operate and enhance the warehouse of which the largest portion was spent on personnel costs and hosting services to store the data. The department currently has 12 full-time positions allocated to the data warehouse and in Fiscal Year 2008-09, it allocated over \$331,000 for hosting services at the state’s Shared Resource Center.¹ (See Appendix A for a breakdown of warehouse expenditures by year.)

Exhibit 1
The Department Has Spent \$20.9 Million for the Development and Operation of the Education Data Warehouse Since 2000

Expenditure	Initial Development (2000-2003)	Ongoing Development and Operations (2004-2009) ¹	Total
Contracted Services	\$6,220,725	\$1,073,785	\$7,294,510
Personnel	2,459,199	4,944,482	7,403,681
Hosting	1,419,206	2,832,404	4,251,610
Expenses	422,220	542,174	964,394
Hardware	282,553	146,295	428,848
Software	216,452	353,493	569,945
Operating Capital Outlay	9,715	1,442	11,157
Total Expenditures	\$11,030,070	\$9,894,075	\$20,924,145

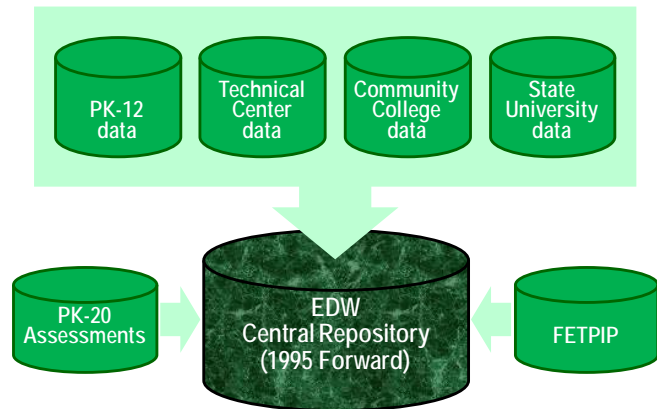
¹ The 2008-09 figures are through March 31, 2009. The initial development phase of the warehouse was completed during the 2003-04 year. The first full year of production was 2004-05.

Source: OPPAGA analysis of Department of Education data.

The Education Data Warehouse primarily operates by storing and retrieving prior-year data rather than collecting and processing current-year data, and does not replace the functions of the department’s primary data systems. To do this, the warehouse links data from 27 statewide data systems. As shown in Exhibit 2, the warehouse integrates and stores selected historical data (from 1995-96 to the present) from multiple database systems. These databases include the source systems for K-12, technical centers, community colleges, and universities as well as the Florida Education and Training Placement Information Program (FETPIP) and the teacher certification database. (See Appendix B for more information.)

¹ In addition to state funds, in 2005-06, the department received a three-year federal grant in the amount of \$1.6 million to expand the types of data incorporated into the warehouse and to enhance data reporting capabilities.

Exhibit 2
The Education Data Warehouse Integrates Historical Data from Multiple Existing Database Systems



Source: Florida Department of Education.

Different entities govern and manage the various database systems. The Florida Board of Education and the Commissioner of Education govern and provide executive leadership over the Education Data Warehouse and the K-12, technical center and community college source database systems. The Division of Accountability, Research, and Measurement within the department is responsible for overseeing and managing these database systems as well as other functions.² The Florida Board of Governors and the Chancellor of the State University System govern and provide executive leadership over the State University Database System.

Findings

The Education Data Warehouse has improved Florida’s ability to track and assess student progress and performance over time, from prekindergarten through graduate school. However, the warehouse has not fully implemented tools to improve stakeholders’ access to data. As a result, the state has not realized the full benefit of its investment. In addition, the department has not sufficiently managed its efforts to improve the warehouse and major source database systems.

² The department also has a separate chief information officer whose responsibilities include providing network connectivity, network security, the department’s email system, computer support, software training programs, and application development.

The Education Data Warehouse has improved the state’s ability to conduct studies of student performance over time

The data warehouse has significantly improved the availability of information for tracking students’ academic progress over time while enhancing procedures for protecting student privacy.

The warehouse provides a more efficient and consistent process for compiling longitudinal student data. The warehouse matches individual student records from the various source database systems to enable educators, researchers, and policymakers to track student outcomes over time. Specifically, the warehouse creates a unique identifier for each student that is used to link and store all of the student’s data in the system. Whenever new data records are loaded into the warehouse, the system performs a matching process to determine if the records match an existing student.³ If there is a 98% certainty that the records match, the new data record is linked to the assigned student identifier; if no match is found, a new student identifier is created. This process builds a longitudinal file for individual students that tracks their demographic, course enrollment, attendance, assessment, financial aid, and program completion information throughout their educational careers.

This matching and tracking capability substantially enhanced the state’s ability to monitor student progress over time. While it was possible to conduct longitudinal analyses across education sectors prior to the warehouse, these studies required extensive efforts to match student records across multiple databases and time periods. The warehouse, by matching and consolidating student data records into one repository, greatly streamlines this process and provides a high assurance that consolidated student records are properly matched and accurate. Because the warehouse stores student records on computer disks, the data is also quickly accessible; in the past researchers often had to wait until data were

³ The records are matched on a combination of Social Security number, gender, race, birth date, and other demographic variables.

loaded from multiple computer tapes that stored information from the various source databases.

The warehouse has established enhanced procedures to protect student privacy. The Education Data Warehouse also has established procedures that have improved the confidentiality of student records. The warehouse assigns a unique identifier for each student, which is used to designate the student's records in lieu of his or her name and Social Security number. The student's confidential information is stored in a separate database and only accessed when data are imported into the warehouse.⁴ In compliance with Family Educational Rights and Privacy Act regulations (FERPA), this enables the warehouse to provide student-level data for longitudinal analysis without disclosing student identifying information.⁵ Before the warehouse was established, the department had to provide personally identifiable information to authorized researchers conducting longitudinal studies in order for the student's data records to be matched over time and across database systems.

The warehouse serves a range of users who report varying degrees of satisfaction. A wide variety of users obtain data from the warehouse, including department employees, external researchers, managers, and policymakers. Users internal to the department generally reported satisfaction, while external users reported mixed satisfaction with the warehouse's services.⁶

Department employees regularly use warehouse data to produce a range of reports on Florida's educational system. These include compiling school grades under the A++ Plan, preparing reports on College Placement Test results, producing the High School Feedback Report, and compiling and reporting civil rights data to the federal government. In addition, academic

researchers use the warehouse to provide data for research studies, and legislative staff use warehouse data to analyze policy proposals being considered by the Legislature.

These users submit data requests to the warehouse, and response times generally depend on the nature of the request and warehouse staff workload. The department reports that, while the data warehouse has 12 FTE, only 3 are dedicated to responding to programming data requests. Relatively straightforward data requests can be filled within a few days, while more complex or lower priority requests may take several months to fill. In some cases, delays are due to inaccurate requestor specifications about the data needed to answer a research question or errors by the department in compiling the requested data. Both of these may result in additional computer programming to produce the needed information.

Both researchers and legislative staff have used source system and warehouse data to help assess and evaluate education policy issues. In general, external researchers reported success in obtaining education data. Legislative staff reported mixed results in obtaining needed data from the department's source systems and/or the warehouse. These staff members often need highly complex data sets from the warehouse and/or point-in-time information from the department's source systems within short time frames to analyze policy proposals being considered by the Legislature. As a result, staff reported that they frequently obtain needed information using personal department contacts who can access and deliver needed data. Staff without department contacts expressed frustration at their occasional inability to obtain needed information in a timely manner.

The warehouse has not fully implemented tools to improve stakeholders' access to education data

While the Education Data Warehouse provides a more efficient system for compiling and storing longitudinal data, it has not fully implemented tools to improve how key users (i.e., state policymakers, program administrators, and

⁴ The process of importing data into the warehouse is referred to as ETL, or Extracting, Transforming, and Loading.

⁵ There are some limited exceptions when data is shared across agencies as governed by federal restrictions.

⁶ To assess user satisfaction with the Education Data Warehouse, we interviewed approximately 45 department data users. We also interviewed a stratified random sample of 26 external users including university and community college staff, legislative staff, and academic researchers.

educators) access the data for analysis and reporting purposes. As a result, the department has not accomplished all of its original goals for the warehouse, the state has not realized the full benefit of its investment in the warehouse, and the purpose for creating the warehouse is not fully realized. This limited data access also contributes to the dissatisfaction some stakeholders expressed with the warehouse.

The department has not fully implemented a key component of the data warehouse. One of the original goals for the Education Data Warehouse was to provide tools that would allow stakeholders to run their own queries of summarized data in a timely, efficient manner. The warehouse was intended to provide advanced reporting and analysis tools—known as business intelligence—that would enable users to use customized drill-down reporting capabilities and subject-specific data marts.⁷ Without these tools, the warehouse provides an abundance of data but limited tools to efficiently access and analyze this information. As a result, access to its data and reports are limited and requires warehouse staff trained in database structure and tools to fill information requests through data extracts. In many cases, this can be a time-consuming process involving several exchanges of communications between warehouse staff and the requester to ensure that the data extracted meets the user’s needs. As part of a recent federal grant award, the department plans to develop an interface “sandbox” that will provide a secure web-based environment that will allow approved, restricted users to extract and manipulate data needed to conduct research. According to the grant application, this self-service function will not be available until the fifth year of the grant in 2013-14.

The department has begun to develop and implement some mechanisms to improve the accessibility of warehouse data. Specifically, the School Grades data mart, the FCAT data mart, and the Performance Profile data mart, which allows

users to view student performance data by legislative district, are available on the department’s website. However, most of the department’s data marts can currently be accessed only by trained warehouse personnel. These data marts allow warehouse personnel to respond more quickly to requests submitted to them on topics for which a data mart exists, but provide no direct end-user access to department staff, the public, or the Legislature. According to department staff, there are plans to deploy data marts for teacher, school safety incidents, and facilities information within the next three months. However, it is unclear whether these data marts will be accompanied by reporting tools that will allow direct access for end users. Other data marts such as those for articulation and FETPIP do not yet have planned completion dates.

The department has not sufficiently managed efforts to improve the warehouse and its source database systems

The department lacks a coordinated approach for improving information collected and maintained by its source database systems and the Education Data Warehouse. In addition, the department has not established adequate planning and oversight processes for some of its major information technology projects. As a result, the department risks duplicating efforts, incurring unnecessary costs, and building database systems that fail to meet the state’s information needs.

The department lacks a coordinated approach for improving the warehouse and its source database systems. Each source system collects, validates, and edits data received from local education entities through a series of iterative steps involving technologies and processes at the Northwest Regional Data Center. Portions of this data are incorporated into the data warehouse. (See Appendix C for more information.) However, the department lacks a comprehensive plan that details how the Education Data Warehouse and its source database systems are to meet the state’s long-term information technology needs. As a result, the warehouse and the source database systems have developed different technology processes and

⁷There are many definitions of "data mart". As described by warehouse personnel, a data mart is a subset of warehouse data relating to a particular topic. The subset is stored in its own database or other file structure external to the warehouse, and is used to provide fast response to queries on that topic.

strategies, which have not been sufficiently coordinated.

For example, over the past two years, the department has been planning to replace aging technologies supporting the K-12 Staff and Student Database system. In addition, the department has separately begun pursuing a vision for expanding and improving the technology for the Education Data Warehouse which includes bringing the source systems into a staging area of the warehouse rather than continuing to have them exist as silos outside the warehouse. Department staff in each of the units indicated that not all parties were 'on board' with the broader vision to bring the former 'silo' systems into the warehouse staging area. While the department reports that the major parties involved are now working together, the lack of initial coordination may have wasted time and resources. To ensure that the merged process represents the most cost-effective and efficient solution for redesigning the K-12 database system, it is important that the department work towards achieving an enterprise-wide approach to IT projects, with a single, integrated, long-term vision.

The department also has not established an effective process for ensuring that its source database systems use common definitions for similar data elements.⁸ These systems are used to support data compilation for many state and federal reports as well as to provide data to the warehouse. The systems often use inconsistent definitions for important data fields. For example, each of the three major source data systems uses different codes for student race. The university system uses federal race codes which include numerous race categories including "non-resident alien", while the K-12 and community college systems each have different codes. As a result, a student's race could be recorded one way while they are in high school, recorded a second way if

they enroll in a community college, and then recorded a third way if they subsequently enroll in a state university. Such differences in data definition can result in inconsistent reporting of key information about students as they progress through Florida's education system. To begin to address this problem, the department recently received a \$2.45-million federal grant to implement a more effective data governance process, which includes developing common data definitions.

The department should develop and adopt a comprehensive plan for the Education Data Warehouse and its source database systems to provide long-term guidance in system development and coordination. This plan should define the role of the warehouse and each of the source database systems, delineate an approach for ensuring that consistent data definitions are used, and establish protocols for ensuring that efforts to improve hardware and software systems are not duplicative and/or incompatible. While current statute does not give the department the authority to require the State University System to use a specific coding scheme, it is in the state's best interest for the department to work cooperatively both internally among its sectors and with the Board of Governors to implement common data definitions where possible to further improve Florida's education data analysis capabilities.

The department needs to strengthen its planning and oversight processes for major information technology projects. The department has undertaken several major projects to improve the warehouse and the source database systems without establishing adequate project management controls and safeguards. Specifically, the department has not established sufficient planning and oversight processes for the continued development of warehouse reporting tools and the K-12 system redesign. Given the importance and magnitude of warehouse development and the K-12 redesign, lack of proper planning could have costly and long-term negative consequences on meeting the state's education information needs.

Proper planning is critical to successfully develop and implement information technology projects. A strong planning process helps ensure that projects

⁸ In 2009, the U.S. Department of Education's Institute of Education Sciences awarded grants to 27 state education agencies to design and implement statewide longitudinal data systems. The Florida Department of Education received a \$2.45 million grant to establish a more effective data governance process which includes developing common definitions across education sectors. In addition, the department will develop a research tool for use by internal and external researchers.

stay within scope, meet planned cost and schedule estimates, and produce desired outcomes when completed. However, the department’s planning processes for the warehouse reporting tools and the K-12 system redesign projects are incomplete. For example, while the department’s plan for the K-12 redesign identifies specific tasks and work to be completed for the project, the plan lacks information about the projects’ scope, objectives, assumptions, risks, and anticipated costs and durations. In addition, the department has not established any formal, written plans to guide its continued development and implementation of warehouse reporting tools. This may have contributed to the department’s delay in implementing the reporting tools. Regardless of whether a major technological endeavor is defined as a ‘project’ or a ‘standard operating procedure’, it is important to define how that endeavor fits into the department’s overall technology development plans.

Information technology projects should also have a project steering committee to help guide the project planning process, monitor project activities, and enhance communication with end users. However, the department has not established formal steering committees for either the warehouse reporting tools or K-12 redesign projects. The lack of a formal steering committee for these major projects increases the risk that projects will not meet the needs of end users and achieve intended project outcomes. It also decreases the chance that stakeholders will buy into and support new functions and applications.

Recommendations

To strengthen the Department of Education’s planning and oversight of its primary database systems, we recommend that it take the actions described below.

- Develop an integrated technology plan for its major education database systems. The plan should detail the department’s long-term vision for the roles and responsibilities of the Education Data Warehouse and the source database systems with regard to data collection, quality assurance, and reporting. In

developing the plan, the department should examine and identify opportunities to streamline overlapping or similar functions and responsibilities across systems.

- Establish a data governance structure or process. The department should regularly convene a group of key program stakeholders to identify and develop common definitions across education sectors.
- Develop formal planning documentation for each project. At a minimum, this documentation should identify the project’s scope, objectives, expected outcomes, assumptions, risks, estimated costs and duration, and specific tasks or work to be completed and the responsible parties for each task.
- Establish project steering committees for major technology projects. The committee’s responsibilities should include assisting in project planning, monitoring project activities, and facilitating communication with project stakeholders. To be most effective, the committee should be composed of representatives from senior management, internal and external users, and technical representatives.

Agency Response

In accordance with the provisions of s. 11.51(5), *Florida Statutes*, a draft of our report was submitted to the Florida Department of Education to review and respond. The Commissioner of Education’s written response has been reprinted herein in Appendix D followed by OPPAGA comments.

Appendix A

The Department Has Spent \$20.9 Million for the Development and Operation of the Education Data Warehouse

As shown in Table A-1, the department's spending patterns for the Education Data Warehouse have shifted over time. From 2000 to 2003, the department spent heavily on contracted services for the initial development and initial implementation of the warehouse. Since that time, the largest portion of warehouse expenses have been for personnel costs and hosting services to store the data.

Table A-1
The Department's Spending Patterns for the Education Data Warehouse Have Shifted Over Time

	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09 (as of March 31)
Contracted Services	\$449,600	\$2,530,058	\$3,195,631	\$ 45,437	\$ 0	\$ 193,286	\$ 464,224	\$ 408,094	\$ 8,180
Personnel	0	582,380	958,433	918,385	1,035,938	1,049,761	982,450	1,047,880	828,454
Hosting	129,679	396,636	283,801	609,091	609,091	635,416	669,247	586,724	331,925
Expenses	18,979	127,919	106,346	168,976	150,918	154,577	135,076	76,083	25,520
Hardware	34,351	\$0	192,042	56,159	8,527	94,019	25,609	0	18,141
Software	31,763	160,784	13,152	10,753	97,520	86,722	147,768	12,920	8,563
Operating Capital Outlay	0	0	9,715	0	0	0	0	1,442	0
Total Expenditures	\$664,372	\$3,797,777	\$4,759,120	\$1,808,801	\$1,901,994	\$2,213,781	\$2,424,374	\$2,133,143	\$1,220,783

Source: OPPAGA analysis of Department of Education data.

Appendix B

Twenty-Seven Statewide Data Systems Feed the Education Data Warehouse

The Education Data Warehouse is a separate repository that integrates and stores selected historical data (from 1995-96 to the present) from multiple, existing database systems. As shown in Table B-1, these databases include the source systems for K-12, community colleges, and universities as well as information from the Florida Education and Training Placement Information Program (FETPIP) and the teacher certification database.

**Table B-1
Twenty-Seven Statewide Data Systems Feed the Education Data Warehouse**

Name of Source System	Description
PK-12 Education Data Systems	
Voluntary Prekindergarten Participants	Data on children who participate in the Voluntary Prekindergarten program. Includes data used to calculate kindergarten readiness rates.
Florida Comprehensive Assessment Test (FCAT)	Florida Comprehensive Assessment Test scores for students in grades 3-10 in reading and mathematics, grades 4, 8, and 10 in writing, and grades 5, 8, and 11 in science.
Bright Futures	Includes both initial eligibility determination for high school seniors as well as continuing eligibility awards for college students.
Course Code Directory (CCD)	A comprehensive information resource that provides general and in-depth information on applicable laws and State Board of Education rules; explanations of requirements and policies pertaining to multiple topics, and details on the K-12 course numbering system.
College Board Data	Advanced Placement and SAT scores for Florida students.
School Grades and AYP	K-12 school grades and AYP (Adequate Yearly Progress) results for the federal No Child Left Behind Act.
Education Facilities Information System (EFIS)	K-12 facilities data, including the Florida Inventory of School Houses.
General Educational Development (GED)	Student scores from the five General Educational Development tests used to receive a high school equivalency diploma.
Florida Kindergarten Readiness Screener (FLKRS)	Student scores on the Florida Kindergarten Readiness Screener (FLKRS), administered to assess the readiness of each child for kindergarten.
PK12 Finance	PK-12 funding and finance data.
PK12 Staff	Demographic information, leave, salary, certification subject and type, professional development
PK12 Student	Student demographic information, as well as course schedule, teacher, exceptional student program, English Language Learner, student transportation, and Title I information.

Name of Source System	Description
Community College and Technical Center Data Systems	
Community College Financial Aid	Student information for students who receive financial aid to attend a public state or community college.
Community College Staff	Staff demographic data as well as academic rank, course load, salary, and benefits.
Community College Student Data Base	Student demographic data as well as high school information, course load, transfer status, entry-level test scores, acceleration mechanisms, program of study, and credit hours earned.
Workforce Development Information System (WDIS)	WDIS is the Workforce Development reporting system and uses five reporting formats. Each format contains specific data elements required for gathering student and teacher data. Includes student demographic and course information.
State University Data Systems	
State University System Financial Aid	Student information for students receiving financial aid to attend a state university.
State University System Staff	State university staff data.
State University System Student	State university student data.
Other Data Systems	
Florida Educational Leadership Examination (FELE)	Florida Educational Leadership Examination registration and scoring data for examinees.
Florida Education and Training Placement Information Program (FETPIP)	Florida Education and Training Placement Information Program contains follow-up data on former students and others, including civilian and federal employment and earnings, continuing education experiences, and military service.
Federal Family Education Loan Program (FFELP)	Data on loan processing, and claims and recovery for federal loans authorized by the Higher Education Act to assist students and their parents obtain help in paying for the cost of higher education, including Subsidized Federal Stafford Loans, Unsubsidized Federal Stafford Loans, Federal PLUS Loans, and Federal Consolidation Loans.
Florida Student Assistance Grant (FSAG)	Florida Student Assistance Grant Program awards is a need-based grant program available to degree-seeking, resident, undergraduate students who demonstrate substantial financial need and are enrolled in participating postsecondary institutions.
Florida Teacher Certification Examinations (FTCE)	Florida Teacher Certification Examinations registration and scoring data for examinees.
Statewide Course Numbering System (SCNS)	A database of postsecondary courses at public vocational-technical centers, community colleges, universities, and participating nonpublic institutions. The assigned numbers describe course content to improve research, assist program planning, and facilitate the transfer of students.
State Student Financial Aid Database (SSFAD)	Student eligibility and award information for postsecondary educational state-funded grants and scholarships such as the Access to Better Learning and Education Grant Program, Critical Teacher Shortage Student Loan Forgiveness Program, Ethics In Business Scholarship Program, First Generation Matching Grant Program, and Florida Resident Access Grant Program.
Teacher Certification	Teacher certification data for those certified in Florida.

Source: OPPAGA analysis.

Appendix C

Florida’s Statewide Education Database Systems Regularly Collect Data from Local Education Entities

The flow of data from school districts, community colleges, and universities to the Department of Education and Board of Governors includes two main processes. The first process collects, validates, and updates data. This process involves collecting data from districts and institutions, performing extensive edit checks, and correcting identified errors. This process is performed using the department's older systems commonly referred to as the source database systems. (See Table C-1.) The department compiles current-year reports almost exclusively from these source systems. The second process integrates data from the source systems into the Education Data Warehouse. Once the first process is complete, selected data elements are moved into the data warehouse where the student records are matched with prior year records and personally identifiable data elements such as student name and Social Security number are replaced with a unique, anonymous student identifier. The data then can be used for longitudinal analyses and reporting. Most multi-year reports on large volumes of data are created from the warehouse.

**Table C-1
The Source Database Systems Have Varying Processes for Collecting, Validating, and Updating Data**

K-12 Student and Staff Database System	Community College and Technical Center Database System	State University Database System
<p>At various predetermined times during the school year, such as the beginning and end of each term, school districts send batch files of student and staff records to the department to be included in this system. These reporting times are commonly referred to as survey periods. The department conducts extensive edit checks of the reported data, and records that do not pass the edits are included on error reports which are sent back to the district for correction. This process is repeated until all data is corrected or the time for corrections expires. Districts usually have approximately a two-week window to initially submit their data, and the update period may last up to several months. For example, for the 2007-08 school year, although the Survey 5 data was due August 2008, updates were allowed for nine months and the data was not final until March 2009. Since modifications are not accepted after the cutoff date, it is possible that some data considered final may not include all corrections. For some types of data, such as staff data, reporting time requirements may vary significantly.</p>	<p>As with the K-12 data system, there are specified survey periods in which community colleges are required to submit data files. Institutions generally have approximately a six-to eight-week period to submit their data in a specified format to the Northwest Regional Data Center (NWRDC). During this time, the institution is able to submit their data, review exception reports, and re-submit data that contained any errors. Files are re-sent to the NWRDC in their entirety after corrections are made. After the close date, no submissions, updates, or modifications by the institutions are accepted. The data is then loaded into a database where it is used for reporting purposes. The Workforce Development Information System (WDIS) data is also collected at NWRDC. After collection, the data is extracted from the mainframe at NWRDC to the department’s servers where validation processing is performed. Following this, the data are used to fulfill operational reporting requirements. Twice a year, the data is extracted and loaded into the data warehouse. The second extract replaces the first.</p>	<p>Standard file submissions from the universities to the Board of Governors are required by term or other period as specified in a calendar on the board’s website. Files having serious errors are returned immediately for correction. Those files passing initial tests are placed in secured locations on an Oracle database, owned by the board and hosted at the NWRDC, where authorized university personnel perform level-two edits. These edit programs are provided by the board and are run by university personnel until errors have been identified and reduced to acceptable levels. Data is then submitted to the board for final editing and acceptance. Once this process has been completed for all universities, the data is released to the universities for various types of analyses. When authorized, university data is pulled into the data warehouse.</p>

Appendix D

FLORIDA DEPARTMENT OF EDUCATION



STATE BOARD OF EDUCATION

T. WILLARD FAIR, *Chairman*

Members

PETER BOULWARE

AKSHAY DESAI

ROBERTO MARTÍNEZ

JOHN R. PADGET

KATHLEEN SHANAHAN

LINDA K. TAYLOR

Eric J. Smith
Commissioner of Education



June 30, 2009

Gary R. VanLandingham, Ph.D.
Director
Office of Program Policy Analysis
and Government Accountability (OPPAGA)
Claude Pepper Building, Room 312
111 West Madison Street
Tallahassee, Florida 32399-0400

Dear Dr. VanLandingham:

Thank you for the opportunity to respond to the draft report entitled *Education Data Warehouse Serves Important Function; Project Planning and Management Need Strengthening*. The Department of Education (DOE) appreciates the opportunity provided for collaboration to ensure the accuracy of the report.

DOE agrees that the Education Data Warehouse was the first of its kind to be developed for education data, is a nationally recognized model for statewide longitudinal data and as stated in OPPAGA's conclusion:

The warehouse provides a more efficient and consistent process for compiling longitudinal student data.

The implementation of the Education Data Warehouse has led to more than \$4 million in federal grants awarded to DOE. These funds, in addition to providing improvements to the Education Data Warehouse, were used to create jobs in Florida. Additionally, because of the Education Data Warehouse's comprehensive data and the ability of stakeholders to access the data, Florida universities have been awarded grants totaling approximately \$11 million.

Dr. Gary R. VanLandingham
June 30, 2009
Page 2

Therefore, the successful implementation of the Education Data Warehouse increases the state's return on investment.

The remainder of this response will address a few remaining concerns regarding the report.

Scope of the Report. The scope of the report mentions that the report is to be both an examination of DOE's Education Data Warehouse (EDW) and an examination of DOE's information technology planning and management efforts for the major source database systems. These are two very different examinations yet the report does not clearly delineate when each area is addressed. Frequently, the two are intermingled so that the results of the project management efforts imply a success or failure on the behalf of the Education Data Warehouse. However, the EDW is not one of DOE's major source database systems. Those systems are the K-12 staff and student databases, the community college staff and student databases, and the workforce education database. The EDW is a longitudinal repository that contains data from these and many other databases, but the EDW is not responsible for the project management or oversight of the source database systems.

This is most noticeable when the K-12 database redesign is used as an example throughout the report of how the EDW has not appropriately applied project management. The K-12 database redesign project has no effect on the efficiency or success of the EDW's functionality. The redesign is an upgrade to one source data system and will not affect the EDW's ability to import the data or to extract the data.

Improving Stakeholders' Access to Data. There are two points to consider when determining the Education Data Warehouse's success regarding the improvement of stakeholders' access to data. First, the EDW has improved access to data when viewing that access long term. What are viewed as simple longitudinal data requests now (those completed in just a few hours) were extraordinarily time consuming and often took weeks to compile prior to the EDW's implementation in 2003. It was an inefficient and oftentimes inaccurate process. But, as with any technology, as user access and ability improves there is more demand for faster capabilities. So, while the EDW has not fully implemented business intelligence tools for public Web access to the data the implementation of the EDW has created a vast improvement over stakeholders' previous ability to access data.

Second, the EDW is aware of the advantages of implementing business intelligence reporting tools, but limited staff and financial resources have prevented the EDW from fully implementing the tools on a wide scale. As mentioned in the report, the EDW has created and implemented several data marts. The report implies that the implementation of the data marts to be used solely by EDW staff is ineffective. However, because these tools have been created for staff, the EDW is able to complete data requests faster than before, thereby improving stakeholder access to data.

Additionally, the business intelligence tools touted by other states regarding their data warehouses are similar to reporting that has been posted on DOE's Web site for years. There is no other model for "easy access" to stakeholders which Florida can replicate. Therefore, to hold Florida to a standard that is far beyond what anyone else in the nation currently has the capability to achieve is unreasonable.

Dr. Gary R. VanLandingham
June 30, 2009
Page 3

Speed of Fulfilling Data Requests. The report mentions that higher priority requests are generally filled within a few days while lower priority requests may take months to complete. DOE would like to point out that high priority requests include requests by DOE staff, the Governor's Office and the Legislature. Lower priority requests include those made by external researchers.

User Satisfaction Interviews. We appreciate the willingness to try to clarify this section (page 4) regarding OPPAGA's interviews with department data users and external users. However, this report is about the Education Data Warehouse. The ability to get data from the source data systems or other databases within the DOE is outside the scope of this report. Yet, it is still unclear whether legislative staff who are displeased with the ability to get data from DOE are referring specifically to the EDW or some other data source.

"[Legislative] Staff without department contacts expressed frustration at their occasional inability to obtain needed information in a timely manner."

Only 26 external users were interviewed and that includes university and community college staff, legislative staff, and academic researchers. DOE would like clarification on how many of the 26 external users were legislative staff, and of those how many were displeased.

DOE has a Governmental Relations Office through which all legislative requests are supposed to filter. The DOE has designed this process so that legislative staff have a single point of contact for all of their requests. DOE needs to ensure that legislative staff are aware of this process so that they can receive their requested information in a timely manner.

The footnote mentions that 45 department data users were interviewed but there is no mention of their satisfaction or dissatisfaction.

Common Definitions. On page 6 of the report there is mention of DOE not establishing a process to ensure the source data systems use common definitions for similar data elements. As mentioned in the report, these systems were developed separately, 20 years ago, under separate governing bodies.

Over the past 2 years DOE's Education Information and Accountability Services, Community College and Technical Center Management Information System, Board of Governors' Information Resource Management, and other data collection systems in DOE have collaborated to ensure that the transition to the new racial/ethnic codes required by the United States Education Department would be seamless and would be useful across systems. This process was implemented successfully and will be the model for one piece of the data governance process being implemented through the recent award of the \$2.45 million Institute of Education Sciences grant.

Assumptions in the Report. A few instances remain on pages 6 and 7 of the report where assumptions are made. These references are misleading because there is no evidence to suggest the causal relationships implied.

Recommendation 1: Develop an integrated technology plan for its major education database systems. DOE has already begun working on an Enterprise Technology Plan. The

Dr. Gary R. VanLandingham
June 30, 2009
Page 4

project charter has been approved and the Scope Statement is being developed. The Enterprise Technology Plan is scheduled to be finalized by December 31, 2009.

Recommendation 2: Establish a data governance structure or process. As mentioned on page 6 of the report, DOE recently was awarded \$2.45 million from the Institute of Education Sciences for a statewide longitudinal data system grant. One piece of the grant proposal is to establish a more effective data governance process. The grant will begin July 1, 2009, at which time a Core Team (including DOE and Board of Governors staff) will convene for one year to determine the roles, responsibilities, and authority of the Data Governance Board. The Data Governance Board is scheduled to convene beginning in July 2010 and continue forward even after the grant has been completed.

Recommendation 3: Develop formal planning documentation for each project. We agree that formal planning documentation should be developed for each project. The Enterprise Technology Plan and the \$2.45 million grant both adhere to project management requirements. The Division of Accountability, Research, and Measurement also now employs two full-time project managers who have been systematically identifying projects and enforcing a typical project management system. However, not all tasks performed by the source data systems are projects. While the tasks may be large and time-consuming, they are often standard operating procedures rather than projects that require project management.

Recommendation 4: Establish project steering committees. We agree with the recommendation of establishing project steering committees for technology projects. However, not all projects warrant a project steering committee. The decision to develop/use a project steering committee should be on a case by case basis and implemented when necessary.

Sincerely,



Dr. Eric J. Smith

EJS/js/br

- c: Sandy Cho
- Linda Champion
- Jeff Sellers
- Ron Lauver

OPPAGA's Comments to the Agency Response

Regarding improving stakeholders' access to data, OPPAGA provides clarification below to the Department of Education's response on page 13.

While the department has implemented some data marts that allow internal, technical staff to respond to data requests faster, it has not implemented the end user access tools that were promised since the inception of the Education Data Warehouse. OPPAGA reported on the system's expectations as expressed by the department in promoting the Education Data Warehouse as it sought continued funding from the Legislature.

OPPAGA supports the Florida Legislature by providing evaluative research and objective analyses to promote government accountability and the efficient and effective use of public resources. This project was conducted in accordance with applicable evaluation standards. Copies of this report in print or alternate accessible format may be obtained by telephone (850/488-0021), by FAX (850/487-9213), in person, or by mail (OPPAGA Report Production, Claude Pepper Building, Room 312, 111 W. Madison St., Tallahassee, FL 32399-1475). Cover photo by Mark Foley.

Florida Monitor: www.oppaga.state.fl.us

Project supervised by Sabrina Hartley (850/487-9232) and Tim Elwell (850/487-9228)

Project conducted by Kathleen Del Monte (850/487-9229), Mike Boland, Rose Cook, Pat Dallet, K. F. Lee, and Brian Underhill

Jane Fletcher, Staff Director, Education Policy Area

Gary R. VanLandingham, Ph.D., OPPAGA Director