Office of Information Technology
Enterprise Data Warehouse

May 1, 2006 to December 14, 2007

Richard L. Fair
State Auditor
The Honorable Jon S. Corzine  
Governor of New Jersey

The Honorable Richard J. Codey  
President of the Senate

The Honorable Joseph J. Roberts, Jr.  
Speaker of the General Assembly

Mr. Albert Porroni  
Executive Director  
Office of Legislative Services

Enclosed is our report on the audit of the Office of Information Technology, Enterprise Data Warehouse for the period of May 1, 2006 to December 14, 2007. If you would like a personal briefing, please call me at (609) 292-3700.

Richard L. Fair  
State Auditor  
March 20, 2008
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>1</td>
</tr>
<tr>
<td>Objectives</td>
<td>1</td>
</tr>
<tr>
<td>Methodology</td>
<td>1</td>
</tr>
<tr>
<td>Conclusions</td>
<td>2</td>
</tr>
<tr>
<td>Findings and Recommendations</td>
<td></td>
</tr>
<tr>
<td>IT Planning</td>
<td>3</td>
</tr>
<tr>
<td>Information Security Program</td>
<td>5</td>
</tr>
<tr>
<td>Data Integrity</td>
<td>9</td>
</tr>
<tr>
<td>Auditee Response</td>
<td>11</td>
</tr>
</tbody>
</table>
Office of Information Technology
Enterprise Data Warehouse

Scope
We have completed an audit of the Office of Information Technology Enterprise Data Warehouse for the period May 1, 2006 to December 14, 2007. Our audit evaluated selected general and application controls related to this system. Our tests of general controls included IT system management planning and development, and information security program planning and management. Our tests of application controls included data processing integrity monitoring.

Objectives
The objective of the audit was to determine the adequacy of selected general and application controls. These controls included policies and procedures to plan, manage, and maintain the system; user authentication to provide system security; and edits and validations to ensure appropriate data input, processing, and output.

This audit was conducted pursuant to the State Auditor's responsibilities as set forth in Article VII, Section 1, Paragraph 6 of the State Constitution and Title 52 of the New Jersey Statutes.

Methodology
Our audit was conducted in accordance with Government Auditing Standards issued by the Comptroller General of the United States. Additional guidance for conduct of the audit was provided by Control Objectives for Information and Related Technology (CobiT) issued by the Information System Audit and Control Foundation.

In preparation for our testing, we studied circular letters promulgated by the Department of the Treasury and policies and guidelines of the agency. Provisions we considered significant were documented and compliance with those requirements was verified. In addition, we obtained and
reviewed software product administration, security, and user manuals. Functions we considered significant were documented and implementation of those features was verified. We also interviewed agency personnel to obtain an understanding of the internal controls.

A nonstatistical sampling approach was used. Our tests of general controls were designed to provide conclusions about the adequacy of those controls in place for IT system management planning and development, and information security program planning and management. Sample transactions were judgmentally selected for testing.

**Conclusions**

Our review disclosed that while an awareness of the need to establish and maintain appropriate controls for the enterprise data warehouse database environment exists, the selected general and application controls for IT system management planning and development, information security program planning and management, and data processing integrity monitoring require improvement. We also noted the Office of Information Technology has spent six years and $8.0 million in the development of an enterprise data warehouse that could provide a reliable source of consistent information for the state as a unified entity, but has been unable to effectively achieve this objective due to a lack of statewide commitment and adequate planning. The decision needs to be made by top levels of state management and strongly communicated to all subsequent levels as to whether a comprehensive enterprise data warehouse is desired. If this type of commitment is not dictated, the necessary agency compliance and coordination may not occur.
A comprehensive enterprise data warehouse has not been developed.

**IT Planning**

An enterprise data warehouse is a database environment dedicated to providing a single, comprehensive view of the enterprise and provides a reliable source of consistent information for financial and strategic decision-making for the enterprise as a whole.

The Office of Information Technology (OIT) has spent six years and $8.0 million in the development of an enterprise data warehouse. They currently have several subject oriented repositories of data for specific sets of users, otherwise known as data marts. The functionality provided within these data marts has been limited to supporting operational reporting.

A lack of planning and monitoring has contributed to this lengthy and incomplete data warehouse development. While OIT has developed a conceptual framework for the management of data known as the Common Information Architecture, it does not maintain a strategic IT plan guiding the development of the enterprise data warehouse architecture. In addition, detailed project plans were not followed. These weaknesses have prevented the accomplishment of the primary objective for implementing the data warehouse architecture, which is data integration.

Strategic IT planning helps to ensure the accomplishment of the business goals for the enterprise. A critical success factor in strategic IT plans is a documented methodology for the IT strategy development which is translated into long-term and short-term plans. Development of the data warehouse architecture began in 2001 with a project to develop a financial data mart for the Office of Management and Budget using an addendum to an existing contract with American Management Systems. At that time, OIT had limited experience and resources available to
enable them to perform this project. This approach ultimately defined the infrastructure software tools and data access tools to be utilized without evaluation of available alternatives. Subsequent projects have either been subsets of the original financial data mart providing a limited single agency view of this data, or have primarily provided agency specific data to that individual agency.

In addition, for project success detailed project plans must be followed. The OIT document titled "Data Warehouse: Development Methodology & Project Management" identifies the primary objective of the data warehouse and data mart project phases as focusing on meeting the state's enterprise reporting requirements. As each project or iteration is completed for a subject area, enterprise-level data warehouse data and data mart specific data should be identified in order to continuously address the statewide enterprise informational requirements. Documentation was provided for a current example project after initial requests for selected completed projects had not been fulfilled. Our review of the documentation disclosed that most of the formal reports and numerous task related deliverables were not available. After inquiry as to the absence of these documents, the conclusion was reached that due to the lack of integration of subject area data the need for the completion of many specified documents associated with this aspect of development was omitted.

Recommendations

We recommend that OIT develop a strategic plan for the enterprise data warehouse architecture which supports the statewide enterprise informational needs associated with the objectives of providing easy accessibility to information and services to the State's residents, businesses, business partners, employees and all levels of government. In addition, we recommend that project plans be followed and required documentation be maintained and approved where applicable, to further the accomplishment
of this development methodology's primary objective of providing enterprise-level data through integration in accordance with enterprise reporting requirements.

Information Security Program

The Information Security Program defines the security goals and standards for the Office of Information Technology (OIT) and has been designed to protect the information resources under its control by formally establishing an ongoing security program. In its efforts to implement this program, a limited number of general policies and procedures were developed. However, due to previous management's inaction, the majority of the information security policies and guidelines identified are currently in a pending status. This has resulted in insufficient policies, procedures, and standards for several critical areas. The following identifies the effects these deficiencies have had on the protection of information under the control of the enterprise data warehouse.

Risk Assessment

An Information Risk Model produced by the Security Oversight Group and a Data Asset Classification System developed by Data Management Services exist but have not been implemented. These initiatives have not been translated into appropriate policies, procedures, and standards to provide guidance in the evaluation of required controls in conjunction with providing assurance as to the confidentiality, integrity, and availability of information.

In accordance with CobiT control objectives, management is responsible for developing a
systematic risk assessment framework. The process should provide for risk assessments at both the enterprise level and system specific level for all projects, and with cross-disciplinary participation. This process should also ensure compliance with external legal, regulatory, and contractual obligations. In addition, a general classification framework should be established with regard to placement of data in information classes (i.e., security categories) as well as allocation of ownership. The access rules for the classes should then be appropriately defined.

**Security Configuration**

The OIT policy, Information Security Vulnerability Management, states "systems shall be hardened in accordance with the applicable OIT Standard prior to release into the operations or development environment". However, the standards referenced have not been formally developed. While system and database administrators have adopted a best practices approach, specific standards should be developed to ensure all vulnerabilities are adequately addressed. Our review of the various servers supporting data warehouse operations identified default groups, users, and services cited in vendor hardening guidelines which should be removed or disabled. In addition, our review of the various databases supporting data warehouse operations identified default users with vendor documented passwords that have not been changed and the availability of unencrypted passwords for privileged user accounts. These items pose vulnerabilities that need to be managed through the development of appropriate controls and standards.

In accordance with CobiT control objectives, IT management should ensure that the set-up of system software to be installed does not jeopardize the security of the underlying data and
programs stored on the system. Attention should be paid to set-up and maintenance of system software parameters.

**Account Administration**

Policies, procedures, and standards for the administration of internal user accounts affording access to open-system based applications have not been developed. Procedures for requesting, issuing, reviewing, and closing of these accounts are not formally maintained and lack consistency. This has resulted in the use of varied user identification formats, ineffective authentication and access mechanisms, and various unnecessary accounts. Our review of user accounts established for the databases used for the data warehouse identified nine non-specific accounts, nine users with multiple accounts, and 14 accounts for users now separated from state service. Our review of the business intelligence software used for the data warehouse identified user accounts without passwords. Generally, these reviews disclosed no password limits, such as expiration times, reuse, or complexity. These issues diminish the overall reliance on user authentication, increasing the risk of unauthorized or inappropriate access.

In accordance with CobiT control objectives, management should establish procedures to ensure timely action relating to requesting, establishing, issuing, suspending, and closing of user accounts. A formal approval procedure outlining the data or system owner granting the access privileges should be included. Controls should be in place to ensure that the identification and access rights of users are established and managed in a specific manner to obtain consistency and efficiency of access control. In addition, procedures should also be in place to keep authentication and access mechanisms effective; for example, regular password changes.
Access Monitoring

The OIT policy, Information Security System Monitoring and User Review, states “system, security, and network administrators shall collect and review system audit logs at prescribed intervals for anomalous activity and possible identification of security-related events”. In addition, it states “computer systems shall have their system audit mechanisms enabled to capture and maintain logs that record security-related activities”. Our review of the servers, databases, and business intelligence software supporting the operations of the data warehouse disclosed no reviews performed and few logs maintained. The lack of monitoring could allow for security related events to occur unnoticed and cause damage to information and information resources.

In accordance with CobiT control objectives, IT security administration should ensure that violation and security activity is logged, reported, reviewed, and appropriately followed up on a regular basis to identify and resolve incidents involving unauthorized activity.

Recommendations

We recommend OIT continue to develop and employ the required policies, procedures, and standards to fully implement their Information Security Program. This includes adopting a risk assessment methodology and data classification system that can be used to identify the necessary controls and standards to ensure achievement of its defined security goals in accordance with the enterprise’s business objectives.
Data Integrity

The State’s Enterprise Data Integration platform, DataStage, supports both high volume batch integration and individual transaction integration in real-time. This Extract, Transform, and Load (ETL) tool is used to move and transform thousands of records in a bulk fashion. An essential component to this process is the capturing of operational meta data produced by DataStage jobs. Operational meta data provides the ability to ensure the accuracy, completeness, and validity of this processed data. The State’s Meta Data Management platform is MetaStage. A major benefit of storing your meta data is the examination of process runs in your warehouse systems. Currently this platform is not operational.

In an effort to manually provide for the assurance of data integrity, Technical Information Documents (TIDs) have been developed to address various issues. A series of these TIDs have been issued to establish standards and procedures for DataStage jobs and sequences and include methods for monitoring job processing. These documents establish standards for job status e-mail notifications and the capturing of job processing statistics. In addition, options regarding the identification of rejected records are also noted. Our review of a 15 day period of provided e-mail notifications disclosed the following.

- Seventy-four percent of the notifications followed the e-mail subject standard for job status. This standard addresses the issue of filtering e-mails to aid in problem identification and resolution.

- Eighty-six percent of the notifications included job processing statistics. This standard provides the method used for logging job run statistics.
Twenty-seven percent of jobs which included job processing statistics identified the capturing of rejected records. The issue of when rejected records are required to be identified has not been clarified.

In addition, a sample of 28 jobs scheduled for execution was tested for receipt of e-mail notifications over this 15 day period. Our review identified 33 percent of the required e-mail notifications were received.

It was also noted during our review that while rows processed may be identified, reconciliations are not routinely performed. In addition, only one job has been identified as including a provision for the reconciliation of processed dollars.

In accordance with CobiT control objectives, processed data should be subjected to a variety of controls to check for accuracy, completeness, and validity. In addition, error handling procedures should be established that enable identification and proper disposition of unprocessed data.

**Recommendations**

We recommend management continue to develop and employ the appropriate technical documentation required to ensure the integrity of processed data. This includes implementation of standards for automating the capture and utilization of operational data in support of process reconciliations and of procedures to assure compliance with standards established for process monitoring.
March 13, 2008

Mr. Stephen M. Eells  
Assistant State Auditor  
Office of Legislative Services  
125 South Warren Street  
Trenton, NJ 08625-0067

RE: The Enterprise Data Warehouse Audit

Dear Mr. Eells:

Recently, the Office of the State Auditor in the Office of Legislative Services completed an audit report entitled “Office of Information Technology Enterprise Data Warehouse”. This report made several conclusions and recommendations. In order to better understand these conclusions and recommendations, it is necessary to consider them in context.

There are many challenges to planning, creating, maintaining, and even auditing an Enterprise Data Warehousing (EDW) environment. As the report points out, this technology domain was new to the State in 2001.

While the audit used general COBIT principles, the specific principles of Control Objectives for Net Centric Technology – Volume Three, Data Warehouse apparently were not used. It was these COBIT principles which became the guide for how the New Jersey Common Information Architecture in general, and its EDW components in particular, were conceived, planned, designed, and implemented.

There are three recommendations made by the audit team which we address below.

First, that OIT develop a strategic plan for the EDW architecture.

At the information architecture level, this strategic plan has existed since the birth of the data warehouse initiative (see attachment “Enterprise Data Warehouse Architecture”, Version 1.00, Prepared January 11, 2001). It is based firmly on COBIT principles for data warehouses, as well as industry best-practices of the time. It is a model for
government information architecture that has been replicated in other jurisdictions. It has always been tied to IT strategic plans for information sharing. What has been missing is a unified IT strategic plan where the information architecture represents substantive, meaningful components with executive sponsorship and funding. As the state moves toward a unified IT Strategic Plan, we will ensure that the EDW plan is consistent.

We do not believe, however, that the production of a project plan for the EDW is feasible as a data warehousing environment is the result of dozens of individual efforts, with a common framework to ensure consistency and reliability. Each of those efforts would have its own project plan with an appropriate level of detail, but collectively they no more make up the project plan for “the data warehouse” than individual building plans make up the project plan for “the city”. In many ways, information architecture is to data what urban planning is to buildings.

The architectures of the following systems have been considered in creating the EDW:

1. The original General Ledger data warehouse,
2. FOCUS, including 204 separate files from 22 mainframe systems,
3. The Medicaid Data Warehouse, and
4. The Taxation Date Warehouse.

In addition, it should be noted that the NJ EDW was created and developed without a dedicated funding source. Instead, it has leveraged a series of five- and low six-figure projects to incrementally grow both its data and capabilities. By using best practices, enterprise-scale tools, and incremental projects, the NJ EDW is creating statewide efficiencies. It is providing reusable data, eliminating redundant interfaces and their associated maintenance and support, and mitigating the costs and risks associated with "big bang" data integration/data warehouse projects.

The EDW’s eight million dollar estimated cost includes the following:

- Initial licensing and cumulative maintenance for the enterprise EDW toolset,
- The pro rata costs and cumulative maintenance for the hosting servers,
- The consulting fees for AMS to develop the first phase of the OMB data warehouse,
- The fulltime employees assigned to the data warehousing and business intelligence center of excellence (4 in FY 2002 growing to 15 in FY 2007), and
- Several fulltime consultants to augment staff shortages in the data warehouse unit.

The tools and hardware as well as the data in the EDW continue to be leveraged. The data warehouse staff continues to take on more projects, providing approximately 1,000 report consumers across seventeen agencies with information from the EDW that may be used for analysis and other management purposes.

Second, that OIT use required security policies, procedures, and standards.

This is sound advice, as the information security (InfoSec) world is constantly evolving with technology and threats. OIT has a multi-tiered approach to security that addresses confidentiality, integrity, and availability – which are, at times, competing interests. It
may be inferred from the audit’s mention of active user accounts for separated employees that this is an area that requires attention. While we will verify that in fact no user account for a separated individual is still available to that individual, we believe the audit comments reflect a miscommunication of the security model in use.

The accounts that were described belonged to database administrators and developers for access to database objects. When individuals separate from state service, several steps are taken.

First and most importantly, their network access is revoked. Without this network access, they do not have the ability to access a database on the state’s network.

Second, as their user ID essentially owns any database objects they have created, we assess whether any of these objects are necessary. If not, we delete the user ID and the objects.

If there are objects we need, we cannot delete the account. Instead, we change the password on the account and make a note that there are objects we will eventually need to migrate. This migration takes place as time is available. Then the account is deleted. Again, once the user separates from state service, their access to the EDW is revoked immediately.

However, to address this recommendation, we will formalize our standard operating procedure in the form of a documented operational requirement that ensures user IDs for anyone separating from service have their password reset if the ID must be maintained, and that if so, an expedited schedule to migrate the objects be identified and executed.

**Third, that management continues to develop and employ the appropriate technical documentation required to ensure the integrity of processed data.**

This is a critical recommendation. We identified this area to the audit team as our most serious concern, and one that required more attention from management and the developers. Over the course of the audit, we have continued to improve the processes that document the loading of data and verify the quality of those loads. We recently put in place an automated process that raises exceptions to both developers and business users based upon data quality and process metrics.

This system replaced a manual system and was developed with feedback from the audit team, and we are indebted to them for their suggestions. Much more still needs to be done. We must convince business users looking to move data into the data warehouse that we must first model that data, define it consistently, develop quality control checks, and implement those in our production process.

Of course, we cannot just dump data into the database and connect reports if we are to maximize the usability of the data and stand behind the quality of the data product. However it is essential that they become a required condition of all future efforts.

**In Conclusion**

We believe the EDW environment has substantiated, substantial value to the State of New Jersey. We believe that the creation of this environment from a standing start in
mid-2001 is a tremendous achievement for the State, especially in consideration of the lack of funding and in comparison to all of the other efforts that have cost more and returned less value.

We know that this environment will continue to grow, and we know that it can be improved, so we will use the audit report as one of the inputs into our continuous improvement process.

Very truly yours,

Adel W. Ebeid
Chief Technology Officer

c: Ronald Stewart, Office of Information Technology
Kenneth Herbst Office of Information Technology
Odysseus Marcopolus Office of Information Technology