

***FSA Integration Partner***  
**United States Department of Education**  
**Federal Student Aid**



**Data Strategy Enterprise-Wide**  
**FSA Data Strategy Framework**  
**123.1.1 Statement of Strategic Data Focus Areas**

***Task Order #123***

**Version 1.1**

April 30, 2003



**Data Strategy Enterprise-Wide  
FSA Data Strategy Framework  
123.1.1 Statement of Strategic Data Focus Areas**

---

**Amendment History**

<b>DATE</b>	<b>SECTION/ PAGE</b>	<b>DESCRIPTION</b>	<b>REQUESTED BY</b>	<b>MADE BY</b>
04/21/03	All	Draft	N/A	N. Brown
04/25/03	3.1	XML, Tech Strategy, CSID Updates	XML, TS, CSID	N. Brown
04/28/03	2.3, 2.4, 3.1.4	FSA Core Team Comments	Core Team	N. Brown
04/30/03	1	Final Revisions	Team Leadership	N. Brown



## **Table of Contents**

<b>1</b>	<b>INTRODUCTION .....</b>	<b>4</b>
1.1	DATA STRATEGY EXECUTIVE SUMMARY .....	4
1.2	WHAT IS AN ENTERPRISE DATA STRATEGY .....	5
1.3	HOW TO CREATE AN ENTERPRISE DATA STRATEGY.....	5
<b>2</b>	<b>THE FSA DATA STRATEGY.....</b>	<b>7</b>
2.1	WHY DOES FSA NEED A DATA STRATEGIES FRAMEWORK?.....	7
2.2	MEETING THE OBJECTIVES THROUGH THE FRAMEWORK.....	8
<b>3</b>	<b>THE FRAMEWORK.....</b>	<b>11</b>
3.1	FSA DATA STRATEGY FRAMEWORK.....	11
3.1.1	<i>Data Access Methods.....</i>	<i>13</i>
3.1.2	<i>Data Standards.....</i>	<i>15</i>
3.1.3	<i>Data Quality.....</i>	<i>16</i>
3.1.4	<i>Data Architecture.....</i>	<i>19</i>
3.1.5	<i>Data Ownership.....</i>	<i>25</i>



## **1 Introduction**

### **1.1 Data Strategy Executive Summary**

The purpose of the FSA Enterprise Data Strategy is to define FSA's enterprise data vision and strategy for how it will combine the tools, techniques and processes, documented in the FSA Data Strategy Framework (Figure 3.1), to handle its enterprise data needs. Specifically, the Strategy will focus on the integration of the following components addressing FSA's major data-related areas:

- FSA Data Strategy Framework
- Technical Strategies
- XML Framework
- Common Identifiers
- Institution Enrollment and Access Management

The Strategy is being crafted in response to Federal Student Aid (FSA) desire to deliver overall improvements in the areas of data quality and data consistency. FSA is also focused on its overall approach for streamlining and integrating systems to ensure that accurate and consistent data is exchanged between its customers, partners, and compliance and oversight organizations. Further, FSA will leverage this data strategy to support the following program-wide business objectives:

- Reduce redundant data storage
- Improve customer service
- Increase accuracy of analytics
- Increase efficiency in data handling
- Reduce costs
- Remove FSA from the GAO high-risk list
- Maintain a clean audit

The purpose of this document is to define these areas and their various components in a broad sense in order to provide the basis for the intended areas of focus for the work to be performed under the FSA Data Strategy task order. These areas will be refined in greater detail by their corresponding teams and combined by the FSA Framework Team into the Data Framework Specification (later deliverable within this task order). That deliverable will provide a vision of the possible future implementation of the tools, techniques and processes listed within this document, forming a strategy for achieving FSA's business objectives.



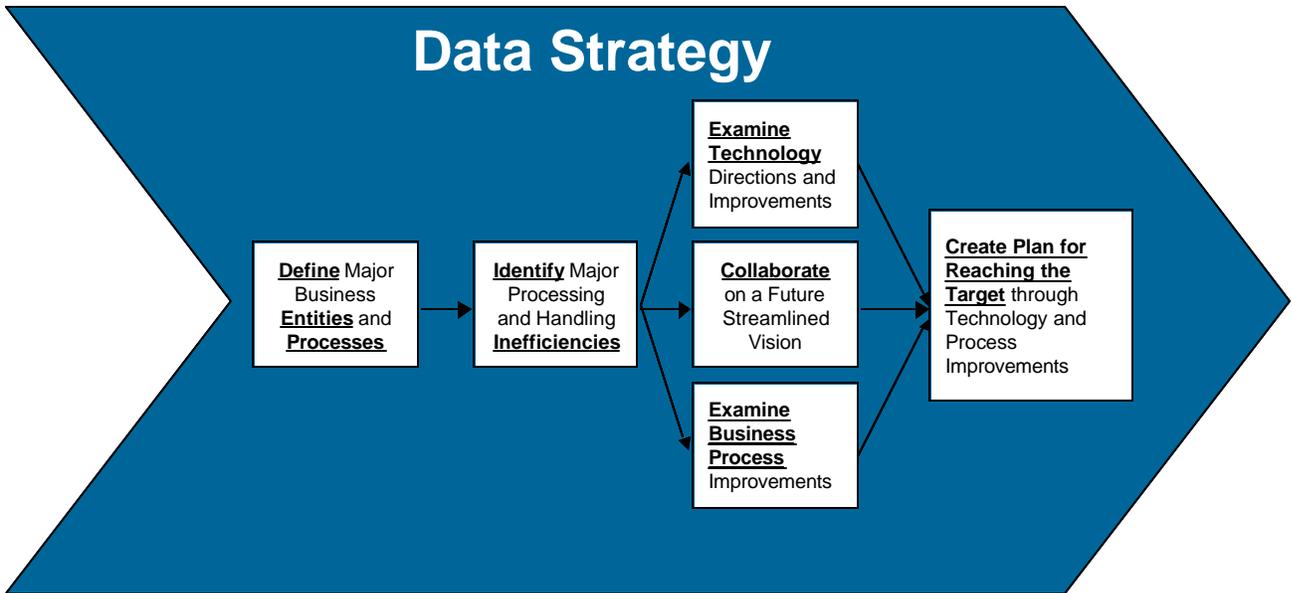
## **1.2 What is an Enterprise Data Strategy**

In order to define an enterprise data strategy a key differentiation must be made between data and information. Data are the bits and pieces of raw facts, the numbers, names, and dates, that when examined in an organized environment and presented in a consistent way provide useful information. Data in an unorganized and inconsistent environment without responsible stewards, is just data. This data provides limited and often misleading information.

This leads us to the basic definition of an enterprise data strategy. It is the development and implementation of a careful plan for the identification, classification, capture, retention, and use of data. When properly created this plan yields an environment that provides useful information regarding the entities that comprise the organization. The overall objective of this project is to complete the development of a plan for an enterprise data strategy for FSA. This strategy, and ultimate solution, will help deliver improved quality and service to customers and enable better management decisions, creating a more responsive and cost efficient organization.

## **1.3 How to create an Enterprise Data Strategy**

To arrive at this strategy it is imperative to understand and document not only the major business entities that an organization uses to process data, but how these entities are handled by the major business processes performed by the organization. Next, an exercise in defining the current processing bottlenecks and inefficiencies needs to be performed. Based on the identification of these issues, the teams within the organization must collaborate on a new vision to handle its business entities and perform its processes through the streamlining of its data handling and processing. This streamlining can be achieved not only through technological advances, but through business process reengineering as well. A sound data strategy will use both of these avenues to achieve success.



**Figure 1.3 Data Strategy Process**

Finally, a strategy is more than just a vision. It is the plan for navigating from the starting point to the finished product. This plan needs to consider technical environments from the current organization and the advancements they are pursuing, and how they can best support the transformation. It needs to consider business processing changes, including when and how customers and entities enter its systems, and how their passage through the systems of the organization can be streamlined to maintain continuity. It needs to consider changing business environments and political impacts to the organization, and how these will affect the needs and uses of the information housed and processed by the organization. All of these aspects are part of the building blocks that make up the FSA Data Strategy Framework.



## **2 The FSA Data Strategy**

### **2.1 Why Does FSA Need a Data Strategies Framework?**

- FSA receives hundreds of thousands of data records from dozens of trading partners every day. In addition, it exchanges and distributes millions of records between its more than 10 internal systems.
- In 2000, FSA began drafting the XML Common Record. The Common Record is the vehicle that standardized the definition and format used to communicate FSA origination and disbursement data universally. This Common Record work indirectly highlighted a major gap in the business processes of FSA; there is no universal definition for other data FSA exchanges externally let alone internally.
- Over the course of the last three years multiple data marts have been created, but there is no integrated plan for the maintenance, stewardship and governance of these marts as they grow and mature. FSA and the Department of Education have aggressive goals to mine data from these various sources for analytical and decision support purposes.
- The birth and growth of these technologies at FSA have created or increased demands on financial and human resources that necessitate the refinement of existing data storage and usage levels to keep costs and budgets in check.
- FSA wants to be removed from the GAO High-Risk List.
- FSA and the Department of Education want to continue to receive clean audit findings.

When considered together, these answers combine to form a business rationale for the creation of an overall plan for how FSA will manage its valuable business data both today and tomorrow – the FSA Data Strategy. These needs led to the creation of Business Objectives in the form of the Fiscal Year (FY) 2003 FSA Annual Performance Plan, which stated that:

- FSA needs a systematic and integrated way of looking at the entirety of its data in order to be able to make informed decisions and improve program integrity.
- FSA can improve program compliance, program monitoring, and program integrity with a consistent and integrated business intelligence infrastructure.
- As a Performance Based Organization, FSA must document its productivity and performance trends. Currently, there is neither an enterprise- wide set of performance metrics nor an integrated reporting mechanism for assessing productivity.
- FSA needs to address weaknesses and conditions to ensure financial integrity, improved customer confidence, security of data and improved data for management decision-making.



**Data Strategy Enterprise-Wide  
FSA Data Strategy Framework  
123.1.1 Statement of Strategic Data Focus Areas**

**2.2 Meeting the Objectives through the Framework**

Senior FSA leadership has created several action items within the plan that are designed to remove FSA from the GAO High-Risk List. The Data Strategy is specifically addressed by action item# 16, and is aimed at ensuring that FSA has relevant, reliable and timely information to effectively manage its programs. This action item identifies the need to define an enterprise-wide data strategy and high-level implementation approach that addresses the business flow of data across the enterprise, architecture, primary ownership, standards, management, access methods, security and quality.

The key action item sub-items addressed by the FSA Data Strategy are listed and mapped to the deliverables within the Data Strategy work as illustrated in the following matrix:

FSA's Risk Plan Action Item			TO 123 Data Strategy Deliverable	
FSA ID No.	Action Item	Map	Del. Num	Name
16	Define an enterprise-wide data strategy and high-level implementation approach that addresses the business flow of data across the enterprise, architecture, primary ownership, standards, management, access methods, and quality.	→	123.1.1	Statement of Strategic Data Focus Areas
			123.1.4	Data Framework Technical Specification
16.1	Identify the strategic focus areas necessary to develop a cohesive enterprise-wide data strategy.	→	123.1.1	Statement of Strategic Data Focus Areas
16.2	Collaborate with all internal Department stakeholders and external stakeholders to identify business needs and requirements with respect to the data provided FSA by others, provided by FSA to others, and managed by FSA.	→	123.1.3	Data Quality Mad Dog Report
16.2.1	Map the current state and future state business flow of data, as applicable, across the enterprise.	→	123.1.2	As-Is System Data Flows
			123.1.4	Data Framework Technical Specification
16.2.2	Develop requirements and initial design for Common Identifiers for School, Students.	→	123.1.22	CSID High-Level Design
			123.1.25	RID High-Level Design

**Table 2.4 TO 123 Deliverable Mapping**



**Data Strategy Enterprise-Wide  
FSA Data Strategy Framework  
123.1.1 Statement of Strategic Data Focus Areas**

FSA's Risk Plan Action Item			TO 123 Data Strategy Deliverable	
FSA ID No.	Action Item	Map	Del. Num	Name
16.3	Determine current data quality and establish target state plan and quality assurance process.	→	123.1.3 123.1.5	Data Quality Mad Dog Report Quality Assurance Strategy and Implementation Plan
16.4	Develop an enterprise-wide extensible markup language (XML) Technical Architecture Framework to enhance data sharing and standardization with our external customers.	→	123.1.13 123.1.14 123.1.15 123.1.16	XML Strategic Assessment and Enterprise Vision XML Technical Reference and Usage Guidelines XML Core Component Dictionaries XML Registry and Repository
16.4.4	Identify initial requirements and initial design for XML Institutional Student Information Report (ISIR).	→	123.1.18a 123.1.18b 123.1.20	Draft XML ISIR Schema Final XML ISIR Schema XML ISIR Technical Reference Support
16.5	Develop integrated Data Warehouse and Data Mart Strategy.	→	123.1.10	Data Storage, Management, and Access Strategy

**Table 2.4 TO 123 Deliverable Mapping (cont'd)**



**Data Strategy Enterprise-Wide  
FSA Data Strategy Framework  
123.1.1 Statement of Strategic Data Focus Areas**

FSA's Risk Plan Action Item			TO 123 Data Strategy Deliverable	
FSA ID No.	Action Item	Map	Del. Num	Name
16.5.3	Develop technical standards and guidelines.	→	123.1.6 123.1.7 123.1.8 123.1.9 123.1.10 123.1.11	Technical Strategies Statement of Strategic Focus Web Usage (Portals) Strategy Web Services Strategy Internal Data Strategy Data Storage, Management, and Access Strategy External Information Access (FSA Gateway) Strategy
16.6	Develop technical standards, conventions, and data management guidelines.	→	123.1.4 123.1.10	Data Framework Technical Specification Data Storage, Management, and Access Strategy
16.8	Develop an enterprise web services/portal strategy.	→	123.1.7	Web Usage (Portals) Strategy

**Table 2.4 TO 123 Deliverable Mapping (cont'd)**



### **3 The Framework**

#### **3.1 FSA Data Strategy Framework**

The FSA Data Framework defines the outline for an integrated FSA Data Strategy. The initial task of the FSA Data Strategy Framework team is to create a Statement of Strategic Focus Areas that further defines the FSA Data Strategy Framework (Figure 3.1). Further, this Statement of Strategic Focus relates this Framework to the business objectives targeted by the 2003 FSA Performance Plan.

The Framework is comprised of five high-level strategic focus areas that make-up the basis of the FSA Data Strategy approach. These areas are:

- Data Access Methods
- Data Standards
- Data Quality
- Data Architecture
- Data Ownership

Through FSA and Integration Partner collaboration, these focus areas have been further defined and refined into the framework diagram illustrated on the following page:



## FSA Data Strategy Framework

<b>Access Methods</b>	Internal Data Exchange <i>(Internal Data Strategy)</i>	External Data Exchange <i>(FSA Gateway Strategy)</i> - SAIG - Web Services	Portals / Websites <i>(Portal Strategy)</i> - Schools - Students - Financial Partners	
	Data Access Services <i>(Web Services Strategy)</i> - EAI - Web Services			
<b>Data Standards</b>	Data Standards - XML - Custom Flat File - EDI		Governance	
			Standards Refinement - Evergreening - Extensions	
<b>Data Quality</b>	Data Correction Services - First time data corrections	Reconciliation Services - Repeatable data consistency checks (Person demo. data) - Data reconciliation scripts	Audit Services - Cross-system financials checks (Aggregated account balance to detailed trans.)	Analytics - Data Mining - Statistical Analysis
<b>Data Architecture</b>	Security - Authentication - Encryption - Authorization - Access - Privacy	Integrated Data Dictionaries	Backup, Restore, and Archiving	XML Vision <i>(XML Strategy)</i>
		Common Identifiers - Common Student ID - Routing ID	Records / Document Management	Core Components / Sector Libraries / Schemas - Refinement - Extensions
		Database / Data Warehouse Strategy <i>(Data Storage, Mgmt., and Access Strategy)</i>		Registry / Repository
				Governance
<b>Data Ownership</b>	Data Owner	Data Owner	Data Owner	Data Owner

\* Representative examples

\*\* Technical Architecture Strategies - In Progress

Last Updated: 03/27/03

**Figure 3.1**



In a practical sense the five strategic focus areas can be thought of using the following analogies:

- Data Access Methods = Format and Transport of Data
- Data Standards = Definition of Data
- Data Quality = Content of Data
- Data Architecture = Management of Data
- Data Ownership = Stewardship of Data

This framework is the living product of extensive thought and discussion around what elements must be defined and refined over time in order to facilitate proper data usage and maintain enterprise data integrity. Details regarding sequence, timing, and implementation will be explored in greater detail within the Data Framework Specification (future deliverable). The outline of the path that FSA will follow is prescribed in the Data Framework diagram illustrated above. First, each focus area is broken down into logical units. These units highlight the stages necessary to build the enterprise data strategy. Each unit is comprised of smaller tasks wherein specific methods and processes will be developed to satisfy the needs of the larger unit. When combined, these units build upon one another to reach consistency within the desired focus area. In turn, these focus areas work together to successfully align the data infrastructure of the FSA Enterprise.

### 3.1.1 Data Access Methods

The top layer of the framework focuses on the access methods that must be defined and standardized so that format and transport of data is consistent throughout the enterprise. The units that comprise the Data Access focus area are:

- Internal and External Data Exchange
- Portals / Website
- Data Access Services

#### *3.1.1.1 Internal and External Data Exchange*

These two items focus on the formatting and transportation of data, or the “what and how”, data flows between systems both internal and external to FSA. Currently data takes on many formats as it is passed between FSA and its trading partners. Some data is exchanged using XML files formats, while others are custom files of varying lengths and data sets. In some cases these custom interfaces even pass excess data between parties.

In addition to content and format variances, data is also exchanged with FSA through various means. Some interfaces are via FTP and secure FTP while others are even exchanged via email and physical magnetic tape. The adoption of data exchange standards and mechanisms for exchanging data in a more uniform way, will lead to fewer unique interfaces and “one-off” exchanges. Defining these standards and choosing specific protocols and transfer methods will achieve reduction in the redundancy of effort and data exchange, paving the way for more



## **Data Strategy Enterprise-Wide FSA Data Strategy Framework 123.1.1 Statement of Strategic Data Focus Areas**

---

efficient transfers. In addition, defining models for internal and external data exchange will set the stage for higher adoption of the FSA EIA architecture while maintaining data integrity.

Information regarding Internal and External Data Exchange will be delivered primarily in Deliverables 123.1.4 Data Framework Specification, 123.1.9 Internal Data Strategy, 123.1.11 External Information Access and 123.1.13 XML Strategic Assessment and Enterprise Vision.

### *3.1.1.2 Portals / Website*

Portal and Website strategies are key technical tools for communicating with FSA's many customers. It is critical that FSA define the standards and processes that will serve as the foundation of the websites and portals that will service these customers so that they are dealt with in a consistent, user-friendly and efficient manner both in terms of user interface as well as maintenance and back-end support. These standards and repeatable processes will increase both customer satisfaction as well as service uptime, while decreasing maintenance costs. While implementing this component of Access Method standardization it is imperative that the existing environments and their architecture be documented, along with suggestions and plans for integration of future services and usage.

Information regarding Portal and Website Strategies will be delivered primarily in Deliverable 123.1.7 Web Usage (Portals) Strategy.

### *3.1.1.3 Data Access Services*

Data access services are the set of tools that enable access to FSA and facilitate the accurate flow of data to its intended destination. Access enablement tools in use today at FSA are EAI Adapters, File Transfer Protocol (FTP) and the Student Aid Internet Gateway (SAIG) mailboxing system. EAI Adapters enable unrestricted sharing of data and business processes throughout the internally networked applications or data sources within an organization. FTP is a common service used by many organizations, including FSA, for the transmission of data between systems. FTP can be used by both internal and external systems for the transmission or receipt of FSA data. The SAIG, or Gateway, is a mechanism for posting and retrieving data traffic through a common location and directing that traffic by assigning addresses of source and destination systems, similar to mailing letters via posted mail. SAIG is also an example of an access method that is applicable to communication both internal and external to FSA.

Another set of access enablers in use at FSA is web portals; the Students, Schools and Financial Partners portals. FSA also wants to explore additional technology tools, such as distributed web services, that can provide data access and transport for "on demand" traffic through these access service points. For example, a web service could be created to execute the processing of an Expected Family Contribution (EFC) using user input through a series of textboxes via the web. These web services can be shared across multiple applications, enabling those applications to provide a user interface consistent with the rest of their application, while providing access to identical, accurate code maintained in a single source to calculate the EFC.



## Data Strategy Enterprise-Wide FSA Data Strategy Framework 123.1.1 Statement of Strategic Data Focus Areas

---

While web services are not access providers in and of themselves, they are key drivers for standardizing and improving customer access services.

Information regarding Data Access Services will be delivered primarily in Deliverables 123.1.4 Data Framework Specification, 123.1.8 Web Services Strategy and 123.1.11 External Information Access.

### 3.1.2 Data Standards

The second layer of the framework focuses on the standards that must be defined so that the meaning and content of data is consistent throughout the enterprise. These data standards are guidelines for how data is defined and represented.

The units that comprise the Data Standards focus area are:

- Standards
- Governance and Refinement

#### 3.1.2.1 Standards

The primary goal for this focus area is to document and provide guidance regarding FSA's data standards and definitions in the following key areas, roughly outlined by their method of exchange:

- XML – eXtensible Markup Language is currently the most widely accepted method for representing data in a portable manner. It is data standard endorsed by the Post-secondary Education Standards Council (PESC). FSA currently uses XML in the Common Origination and Disbursement (COD) Common Record and the XML ISIR Schema (currently in draft version).
- Custom Flat Files – Custom flat files are used throughout the enterprise both internally and externally. These files do not currently have a governing data standard, by which they can be modeled. Definition of data within these files is limited in scope to only those parties on either “end” of the interface, sender or receiver.

Unfortunately there are currently no enterprise-wide standards for how data is defined or represented. Certain blocks of information, for instance those within the COD Common Record, have agreed upon community definitions. These fields represent only a small percentage of data exchanged by and with FSA. Establishing standards for data passed within XML files as well as custom flat files will allow for greater reusability, and will ease the eventual transition from custom flat files to XML. These standards will also create a framework for discussions with PESC regarding how best to govern their usage.

Information regarding Data Standards will be delivered primarily in Deliverables 123.1.4 Data Framework Specification, 123.1.9 Internal Data Strategy and 123.1.15 XML Core Component Dictionaries.



### *3.1.2.2 Governance and Refinement*

It is also important that FSA's Data Standards are consistent with the community (e.g., PESC) and its trading partners, and, where applicable, other agencies and offices within the Federal Government. This is particularly relevant to the activities being performed around aligning data classification and grouping with those of the Department of Education. Creating formal and applicable data standards will create symmetry and reusability for FSA and its customers. These standards are the building blocks that are required to create a data reconciliation and clean up plan as well as sustainable reconciliation services. Further, these standards will create an environment in which decision support analytics are not only possible, but also accurate.

The primary focus of the data standards segment of the FSA Data Strategy work will be to first define which of FSA's current data elements are exchanged in one format or another, then to create a definition of these elements for adoption by PESC and the community. The XML specific data standards will be focused on by the XML Framework team and in concert with the larger Data Strategy Team.

Once defined, a continual process of evolution and refinement is inevitable within the dynamic community of student loan information. Therefore, a plan for how FSA will participate in the process of this refinement will be created so that it plays a planned and repeatable role in this data evolution.

Information regarding Governance and Refinement will be delivered primarily in Deliverables 123.1.5 Quality Assurance Strategy and Implementation Plan and 123.1.13 XML Strategic Assessment and Enterprise Vision..

### *3.1.3 Data Quality*

The third and arguably most desirable layer of the framework focuses on the quality of the data content housed and used by FSA. Data Quality focuses on the accuracy and consistency of the data content throughout the enterprise. The units that comprise the Data Quality focus area are:

- Data Correction Services
- Reconciliation Services
- Audit Services
- Analytics

These units represent a maturity of data quality for the organization. Data Correction will provide a foundation for implementing and automating Reconciliation services. Once a process is in place to keep data in synch, the creation and execution of independent Audits by both internal and external parties will be enabled. Ultimately, these three levels of maturity will combine to create an environment where high quality, reliable Analytics are supported and achievable.



## **Data Strategy Enterprise-Wide FSA Data Strategy Framework 123.1.1 Statement of Strategic Data Focus Areas**

---

It is important to note that while the Data Strategy effort will document the data quality findings through the Mad Dog effort as well as a Quality Assurance and Implementation Plan, the Data Strategy effort will not be executing the quality measures outlined in this section.

### *3.1.3.1 Data Correction Services*

Data correction is the first step in achieving improvements in overall data quality for the enterprise. It is comprised of both initial data edit checks of the large volume of data that is submitted to FSA from external parties as well as the initial efforts to correct data within a system using clean-up scripts and validation rules. In order to take this initial step, it is imperative that not only are data standards for data exchange, edit checks, and usage created and published for use, but that these guidelines are adhered to by all parties. Data correction is not a one-time accomplishment or goal. Data Correction Services are the mechanism both for making initial corrections and ensuring that data remains correct until a series of processes, both automated and manual, are adopted to keep data aligned across the enterprise.

Data Correction is achievable by following a well defined path outlined by internal data standards, common data elements mapping, common definition of exchanged data, and common data pass through points. Knowing who owns data, who has copies of it, and who disseminates all support the goal of Data Correction Services.

Information regarding Data Correction Services will be delivered primarily in Deliverables 123.1.4 Data Framework Specification and 123.1.5 Quality Assurance Strategy and Implementation Plan.

### *3.1.3.2 Reconciliation Services*

Reconciliation Services will be built on the Data Correction services model and will be used to facilitate repeatable data consistency checks. Reconciliation services prescribe a method to maintain FSA's data so that information remains correct and synchronized across systems. For example, systems using a Person's demographic data will have the most up-to-date version even if the person's information has changed in the duration of the Financial Aid Lifecycle. This will be achieved by defining actions that should trigger a cascade of updates when key core component data is updated at various points throughout the lifecycle.

Reconciliation Services may include reconciliation scripts, for instance a "data health meter" or a mechanism that can examine data and determine when updates to stale or affected data should be made. It will establish guidelines for how these scripts will be created as well as provide recommendation regarding when and how often they are administered. It is also possible that these services may take advantage of new web services technology or enhanced EAI services to automate reconciliation in real-time.

Information regarding Reconciliation Services will be delivered primarily in Deliverables 123.1.4 Data Framework Specification and 123.1.5 Quality Assurance Strategy and Implementation Plan.



### *3.1.3.3 Audit Services*

Audit Services will enable cross-system financial checks to validate that financial data aligns across FSA systems. An example of a cross-system check is confirming that FMS issued accurate payments based on detailed data from NSLDS or another discrete system. Using audit services, financial information can be checked at different levels of summarization, from an aggregated account balance to individual detailed transactions.

Implementing audit services at FSA is highly anticipated by the Department of Education and the Office of Management and Budget. Audit services will compliment existing FSA efforts to remain off of the GAO High-Risk List. While Audit Services cannot be implemented immediately, steps in that direction will be taken in the Data Correction and Reconciliation Services work. Additional changes in regulation and data storage will also be required to realize the implementation of Audit Services. This work will examine the substance and sequencing of those changes to enable a more sound and accountable financial picture for FSA.

Information regarding Audit Services will be delivered primarily in Deliverables 123.1.4 Data Framework Specification and 123.1.5 Quality Assurance Strategy and Implementation Plan.

### *3.1.3.4 Analytics*

Analytics, within the framework, encompasses the ability to make informed business decisions based on accurate, consistent data. Analytics includes the use of data mining and statistical analysis tools and procedures. Some analytics abilities exist within FSA today within the following data sources:

- Credit Management Data Mart (CMDM)
- Financial Partners Data Mart (FP Data Mart)
- Delinquent Loans Data Mart (DLDM)
- eCB Data Mart
- eZ Audit Data Mart
- National Student Loan Data System (NSLDS)

Although there are pockets of information and tools for analytical analysis within FSA, there is no enterprise-wide strategy for organizing data for analytics. For instance, NSLDS contains detailed loan and person information about nearly all Title IV loans and enables users to execute or create historical queries to use for analytical purposes, but it is not a formal data warehouse or data mart. In fact, NSLDS is a transactional system performing some warehouse/mart duties. However, it is still not a widely embraced enterprise-wide solution.

In order to realize the true value of analytics, the results must be based on data that is both accurate and consistent. The data should also be at the same level of detail and mined from the same point in time. Analytics plays a large part in the needs analysis portion of the Data Warehouse Technical Strategies initiative and will be coordinated closely with the FSA Data



Framework work. Implementation of enterprise analytics is highly dependent upon the successful implementation of the other preceding data quality initiatives.

Information regarding Analytics will be delivered primarily in Deliverables 123.1.4 Data Framework Specification, 123.1.10 Data Storage, Management and Access Strategy and 123.1.12 Technology Vision and Strategic Plan.

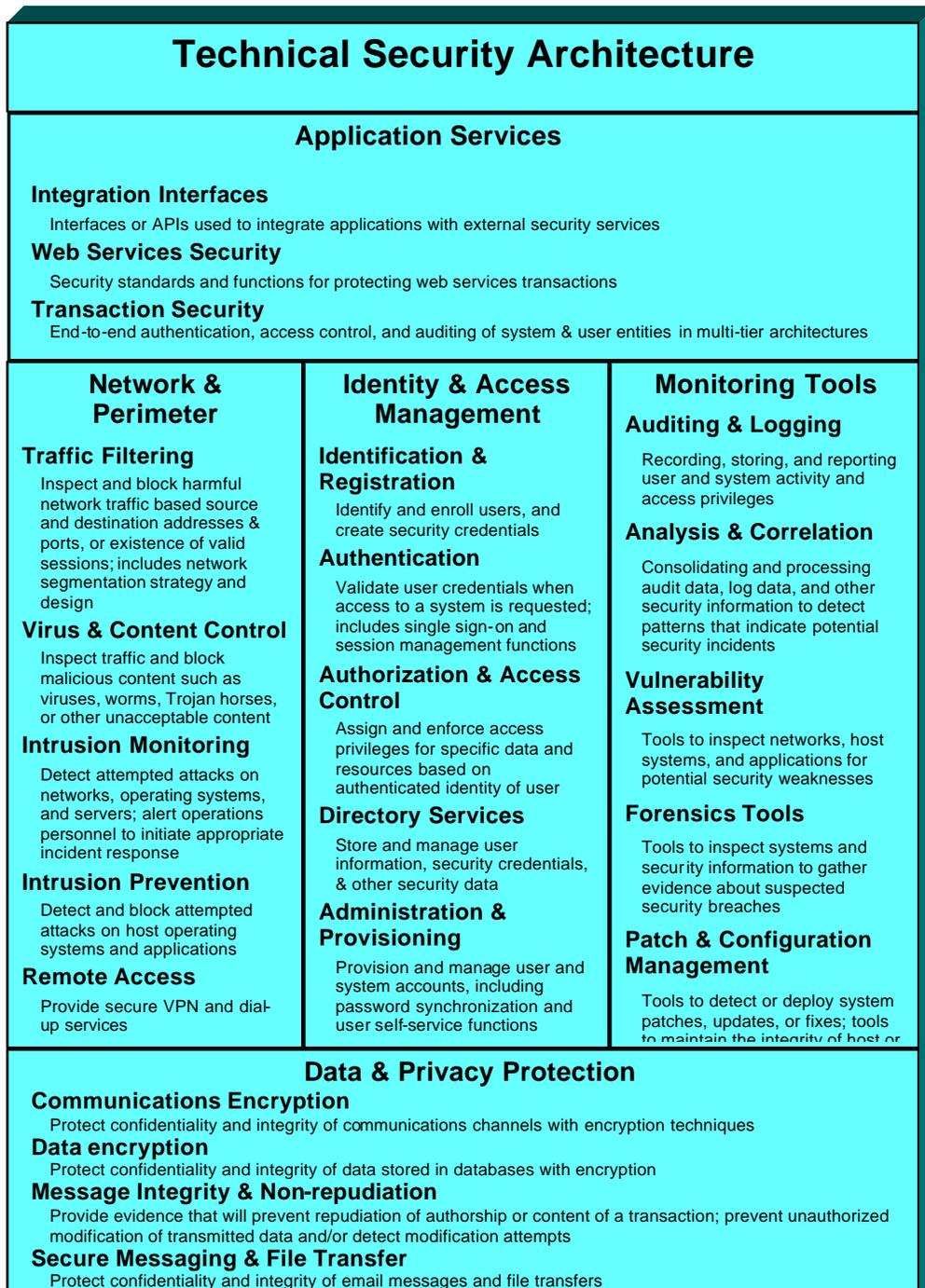
### 3.1.4 Data Architecture

The fourth layer of the framework focuses on the physical architecture, data organization approach models and software tools that FSA uses to manage its data. These architecture components, while technical and tactical in nature, are essential to enabling the functional use, protection, and interpretation of data throughout the enterprise. The units that comprise the Data Architecture focus can really be broken into two subgroups:

- Physical Architecture:
  - Security
- Data Organization and Management Architecture
  - Integrated Data Dictionaries
  - Common Identifiers
  - Database / Data Warehouse Strategy
  - Backup, Restore, and Archiving
  - Records / Documents Management
  - XML Vision
  - Core Components / Sector Libraries / Schemas
  - Registry / Repository
  - Governance

#### 3.1.4.1 Security

Security includes Authentication, Encryption, Authorization, Access, and Privacy of data within the FSA enterprise. It is a critical part of the data framework and will be addressed by a separate, closely integrated initiative that is focused on designing a technical security architecture tailored specifically for FSA. This architecture is based on the following generic technical security framework (Figure 3.1.4.1), and will be augmented through FSA and Integration Partner collaboration to design a strong Security Architecture for FSA.



**Figure 3.1.4.1 Technical Security Architecture**



Information regarding Security will be delivered primarily in deliverables associated with Task Order 124 Security and Privacy Architecture Framework. However, integration of this security framework will be documented in deliverables 123.1.4 Data Framework Specification and 123.1.12 Technology Vision and Strategic Plan.

#### *3.1.4.2 Integrated Data Dictionaries*

The concept of an Integrated Data Dictionary would ultimately include all enterprise data: data that is shared between systems in an XML format, data that is shared between systems in a non-XML format, and data that is exclusive to a system.

While data dictionaries for some individual FSA systems currently exist, the value of creating an enterprise data dictionary for all FSA Enterprise data, especially data that is exclusive to a single system, is somewhat unnecessary in order to arrive at an enterprise-wide strategy. The data that is necessary to codify and organize in a dictionary is that data which is exchanged or shared, and therefore needs to be defined for common understanding and usage. The task of defining and codifying data that is not exchanged has little value across systems during the data strategy definition, and would require a great deal of maintenance and upkeep given the number of proprietary and COTS systems in use at FSA.

The primary focus of the Data Strategy, as it relates to the creation of an Integrated Data Dictionary, will be defining data objects shared between systems. The common data definitions established in the Integrated Data Dictionary and through the XML Frameworks will be the building blocks for the entire Data Framework.

An effort is currently underway with the Department of Education to map all of the data elements in use by FSA to a single classification system. The Data Strategy effort will make every effort to utilize the standards in use for this ongoing effort. However, it is important to define the difference between these efforts. The Data Strategy effort is focusing only on those elements being exchanged while the larger Department effort is targeting all data in use at the Department and FSA.

Information regarding Integrated Data Dictionaries will be delivered primarily in Deliverable 123.1.4 Data Framework Specification, 123.1.9 Internal Data Strategy and 123.1.10 Data Storage, Management and Access Strategy.

#### *3.1.4.3 Common Identifiers*

Common Identifiers are a strategic component of the overall FSA Data Strategy. The Common Identifiers will help FSA view data across the student aid lifecycle, by isolating specific identifiers and rules that systems can use to identify a single student or institution. The Common Identifiers work is the crafting of a high level design and implementation analysis for enterprise-wide student and school identifiers.



Improving these identifiers through a standalone initiative has a limited affect on the overall quality and strategy for FSA's data. They will prove infinitely more useful in conjunction with other pieces of the framework. For example, as overall data quality is improved, data will be more readily searchable and sort-able through common, unique identifiers, drastically improving the usefulness and accuracy of analytical queries and reports.

#### **3.1.4.3.1 CSID**

The lack of an enterprise-wide student (person) ID inhibits FSA's ability to view data about a customer across the phases of the lifecycle. This deficiency on standards enables identification errors. For example, unique customer records can be inappropriately merged creating privacy concerns, and customers' records are not always linked appropriately preventing FSA from viewing data about specific customers across all phases of the life cycle. The CSID initiative seeks to establish a simple framework by which FSA and Trading Partners can identify applicants and borrowers across lifecycle phases and across systems that support these customers.

A great deal of discussion and progress has already been made regarding this initiative. This work is being leveraged as the starting point from which to build a consensus around the specific data elements and processes that will enable a common identifier for a student across their life cycle.

Information regarding CSID will be delivered primarily in Deliverables 123.1.22 CSID High Level Design and 123.1.23 CSID Implementation Strategy.

#### **3.1.4.3.2 Routing ID**

Much like the CSID, the lack of an enterprise wide identifier for a school produces inefficiencies and inconsistencies around data handling for schools. The Routing ID will facilitate the reconciliation of school entities across FSA's legacy systems, improve program integrity through enterprise entity identification, and provide a consistent means of school entity identification across FSA's enterprise.

This initiative is tightly integrated with the need to streamline and enhance the point of entry for the school into the FSA Aid process. The entry point, where schools first apply and are processed for involvement in the various FSA Aid programs, is part of the Enrollment and Access Management process. In addition to this initial entry point, the maintenance of service access and users is also part of this process. The Routing ID is pivotal in minimizing the bureaucracy and increasing the efficiency associated with this critical FSA trading partner touch-point.

Also similar to the CSID initiative, a good deal of work has been performed to date regarding the Routing ID. This work is also being leveraged to jump-start the design and options analysis that will lead to usage of a common Routing ID.



## **Data Strategy Enterprise-Wide FSA Data Strategy Framework 123.1.1 Statement of Strategic Data Focus Areas**

---

Information regarding Routing ID will be delivered primarily in Deliverables 123.1.24 RID Implementation Options Analysis, 123.1.25 RID High Level Design, 123.1.26 Enrollment Business Objectives and High Level Requirements and 123.1.27 Access Management Business Objectives and High Level Requirements.

### *3.1.4.4 Database/Data Warehouse Strategy*

Database/Data Warehouse Strategy is the process of creating a high level plan for FSA data storage through various data housing tools and data organization approaches. This component of the Data Strategy is vital to realizing FSA's business goals and objectives. This is the element of the strategy through which the careful organization and synthesis of the data that FSA collects and stores will transform its data into valuable information for its customers and users.

Currently, FSA has multiple data marts, all housing and "slicing" data in a specialized way for its customer set. However, these marts exist independent of one another, sourced from various systems, and in some cases service similar needs for different customers. Another sizable data store is NSLDS, which performs business tasks related to multiple systems and life cycle phases. At the same time, it acts as a quasi-data mart.

These facts point to the need for an enterprise wide plan for how data is stored, processed and analyzed. The Database/Data Warehouse Strategy is closely tied to the Analytics piece of the data framework and will be critical to developing solutions for how to best meet these analytical and research needs.

Information regarding Database / Data Warehouse Strategy will be delivered primarily in Deliverables 123.1.4 Data Framework Specification, 123.1.9 Internal Data Strategy and 123.1.10 Data Storage, Management and Access Strategy.

### *3.1.4.5 Backup, Restore and Archiving*

Backup, Restore and Archiving involves creating a strategy to effectively administer data backup, data restore, and data archiving. Currently individual FSA systems backup, restore and archive data, however no enterprise-wide guidelines exist. Determining what data should be backed up, and how it should be stored and restored, relate to business continuity planning for the enterprise. Backing up data efficiently ensures critical business information is not lost. Restoring data effectively allows for backed up data to be made available when the business need arises. Proper archiving can lead to reduction in costs and increased performance on databases by moving inactive or historical data to longer term, sometimes "near-real-time" locations.

Information regarding Backup, Restore and Archiving will be delivered primarily in Deliverables 123.1.9 Internal Data Strategy and 123.1.10 Data Storage, Management and Access Strategy.



#### *3.1.4.6 Records and Document Management*

Records and Document Management encompasses how records and documents are accessed and controlled throughout FSA. Specifically, document management involves delivering controlled access, storage, and publication of large volumes of electronic and paper-based content. Document Management would also include electronic processing of paper documents through imaging. Imaging is the process of electronically capturing and reproducing paper-based information to support specific business tasks. Records Management addresses the legal aspects of record retention and compliance with Freedom of Information Act (FOIA) regulations. Records and Document Management guidelines will define how FSA can maintain compliance with these rules in an integrated and enterprise manner.

Information regarding Records and Document Management will be delivered primarily in Deliverables 123.1.9 Internal Data Strategy and 123.1.10 Data Storage, Management and Access Strategy.

#### *3.1.4.7 XML Vision*

XML Vision and Methodology will define the standards and processes for FSA's use of XML across the enterprise. The Vision and Methodology document includes:

- Methodology – The methodology will define how FSA builds-out its XML infrastructure and capability. It will incorporate proven practices from industry case studies.
- Governance – The Governance Model will help FSA consistently use XML across the enterprise; and will address issues such as leadership boards, stakeholder committees, responsibilities, policies and procedures, as they are applicable.
- Processes – Since the financial aid community (e.g., PESC), will own the standards, the processes will need to define how the XML standards are reviewed and approved, as well as the ongoing processes to support changes and updates to the standards.

XML is currently used for data representation within a limited portion of FSA. This representation is in the COD Common Record. It defines and captures the core components exchanged during the loan origination and disbursement processing performed by FSA through COD with participating schools. However, as information needs evolve and grow, it is likely that XML usage and advantages will drive more systems and trading partners to use it in the future.

Information regarding the XML Vision will be delivered primarily in Deliverable 123.1.13 XML Strategic Assessment and Enterprise Vision.

#### *3.1.4.8 Core Components, Sector Libraries, and Schemas*

Core Components, Sector Libraries, and Schemas are the technical components of the XML Framework. The Core Component Dictionaries are the lowest logical groupings of data. They will establish enterprise definitions for the baseline core components used across the student



financial aid lifecycle by FSA's systems. The Core Component Dictionary will aid in the establishment of the Integrated Data Dictionary.

The next higher level of data grouping is the Sector Library. Sector Libraries establish a business domain or context for the XML components. Finally, these libraries are grouped into XML Schemas. XML Schemas define the format of XML documents and will be consistent with the Data Standards defined by the Data Framework team. This process will be the basis for standardizing data exchanged through XML based interfaces across the FSA Enterprise. Those interfaces and exchanges that do not use an XML format will also utilize a similar framework for defining and standardizing data. However, the ultimate interfaces will not be exchanged using an XML-based file.

Information regarding Core Components, Sector Libraries, and Schemas will be delivered primarily in Deliverables 123.1.14 Technical Reference and Usage Guidelines and 123.1.15 XML Core Component Dictionaries.

#### *3.1.4.9 Registry / Repository*

The XML Registry / Repository is the storage mechanism for the technical components of the XML framework. Definitions for the elements, entities and blocks that comprise the Core Components Dictionary, Sector Libraries, and XML Schemas are all stored in the XML Registry / Repository. Essentially, the XML Registry and Repository is a tool that will provide FSA the capability to manage its XML components and documentation. The XML Registry and Repository will enable FSA users to view and download XML relevant documents.

Information regarding XML Registry / Repository will be delivered primarily in Deliverable 123.1.16 XML Registry and Repository.

#### **3.1.5 Data Ownership**

The fifth and final layer of the framework focuses on the ownership of the data housed and used by FSA. Data Ownership refers to identifying business process and data owners who will be the stewards of that data throughout its existence within FSA's systems. More specifically, it is the process of identifying the systems that will be the caretakers of data during specific life cycles as data entities enter, progress and leave FSA's business processes.

The following issues will be addressed regarding data ownership:

- The source or originating system for a data component
- Where and when partial or full copies of data components exist
- Who passes the data component to whom and for what purpose
- Who stores the data component for future use

Addressing and mapping out these issues in the context of the student aid lifecycle and trading partner lifecycle, by major business entity, will enable FSA to define a plan for the stewardship of data components. These major business entities include: Person (Borrower / Applicant), Aid



**Data Strategy Enterprise-Wide  
FSA Data Strategy Framework  
123.1.1 Statement of Strategic Data Focus Areas**

---

(Loan or Grant), School and Financial Partner. This mapping is a vital component for all data quality initiatives, especially the Database/Data Warehouse strategy and the Data Quality Mad Dog effort.

Information regarding Data Ownership will be delivered primarily in Deliverables 123.1.2 As-Is System Data Flows, 123.1.3 Data Quality Mad Dog Report, 123.1.4 Data Framework Specification and 123.1.5 Quality Assurance Strategy and Implementation Plan.