Medicaid Enterprise System (MES) Procurement Project

Strategic Enterprise Advisory Services (SEAS)

T-1: Data Management Strategy

May 30, 2018

Created By: The SEAS Vendor

Submitted To: AHCA MES Project Management

Revision History

DATE	VERSION	DESCRIPTION	AUTHOR
04/09/2018	001	T-1: Data Management Strategy Initial Draft Version (Entry)	Steve Quante, Paul Moore
05/14/2018	002	T-1: Data Management Strategy Revised with Response to Agency Review Comments	Steve Quante, Jessica Rein, Paul Moore
5/25/2018	003	T-1: Data Management Strategy Revised with Response to Agency Review Comments	Steve Quante, Paul Moore
5/30/2018	100	T-1: Data Management Strategy Final Draft	Sean Gibbs

Modifications to the approved baseline version (100) of this artifact must be made in accordance with the Change Control process that is part of the Scope Management Plan.

Quality Review History

DATE	REVIEWER	COMMENTS
04/08/2018	Sean Gibbs	QA Submission Review
05/14/2018	Courtney Sullivan	QA Resubmission Review
5/27/2018	Sean Gibbs	QA Submission Review

Table of Contents

Table of Exhibits

Table of Strategic Topics

SECTION 1   INTRODUCTION

1.1 BACKGROUND

The Florida Agency for Health Care Administration (Agency) is preparing for the changing landscape of health care administration and increased use of the Centers for Medicare and Medicaid Services (CMS) Medicaid Information Technology Architecture (MITA) to improve the administration and operation of the Florida Medicaid Enterprise. The current Florida Medicaid Enterprise includes services, business processes, data management and processes, technical processes within the Agency, and interconnections and touch points with systems that reside outside the Agency necessary for administration of the Florida Medicaid program. The current Florida Medicaid Enterprise System (MES) includes the Florida Medicaid Management Information System (FMMIS), Decision Support System (DSS) and other systems operated by different vendors. These systems in the MES interface through the point-to-point exchange of data files using Secured File Transfer Protocol. These interfaces become more complex and costly as the number of systems and applications increase. The future of the Florida Medicaid Enterprise integration is to allow Florida Medicaid to secure services that can interoperate and communicate without relying on a common platform or technology. Connecting services and infrastructures and developing integration standards are the next steps for advancing the MES level of MITA maturity and system modularity modernization.

CMS released the Medicaid Program Final Rule: Mechanized Claims Processing and Information Retrieval Systems in December 2015. This final rule modifies regulations pertaining to 42 Code of Federal Regulations (CFR) 433 and 45 CFR 95.6111, effective January 1, 2016. Among other changes, this final rule supports increased use of the MITA Framework. MITA is a CMS initiative that fosters an integrated business and information technology (IT) transformation across the Medicaid enterprise to improve the administration and operation of the Medicaid program. The Agency documents its high-level plans to increase service interoperability and advance the maturity of the MES in accordance with the MITA Framework in the Florida MES Procurement Strategy document.

1.2 PURPOSE

The purpose of the MES T-1 Data Management Strategy (MES DMS) is to develop and articulate the Agency's Data Management Strategy. The MES DMS aligns with the MITA 3.0 Part II Information Architecture - Chapter 2 Data Management Strategy (MITA DMS) while accounting for unique Agency requirements. The MES DMS is the product of current state discovery, stakeholder input, strategic analysis, program strategy and direction about techniques and priorities to support overall improvement of Medicaid program outcomes.

The MES DMS document may contain links to later versions of documents and diagrams, referenced within the following sections, that reside in the MES Projects Repository.

The MES DMS communicates MES data management direction and strategy to interested stakeholders including Agency for Health Care Administration (Agency) technology leadership, executives, CMS, existing and potential MES Project Vendors, MES stakeholder agencies (e.g. Department of Children and Families, Department of Health, Agency for Persons with Disabilities, Department of Elder Affairs, etc.), and other state Medicaid programs.

1.3 SCOPE STATEMENT

This iteration of the deliverable provides the foundational direction for MES data management focusing on the technologies, processes, and tools needed to implement the MITA 3.0 Data Management Strategy. This iteration enables the foundational capabilities of Integration Services Integration Platform (ISIP), Enterprise Data Warehouse (EDW), and modular capability implementation. This document is not a detailed implementation manual but provides the Agency context, aligned with MITA, required for planning purposes.

As per MITA guidance, the MES DMS will communicate strategy and direction for the following topics:

- Data Management and Integrity Approach
- Enterprise Data Management and Data Stewardship
- Common Data Architecture
- Enterprise Modeling
- Enterprise Metadata Repository
- Data Sharing Architecture
- MES Specific Additions

1.4 GOALS AND OBJECTIVES

The Goals and Objectives of this document are:

- Goal 1 – Establish the MITA compliant Florida Medicaid Data Management Strategy
  - › Objective 1 – Define and document the core Data Management Strategy areas for the Agency that align to the MITA standard as described in Section 1.3 Scope Statement
  - › Objective 2 – Provide key strategic data management guidance and reference for future procurements as part of the Agency's modular implementation approach
- Goal 2 – Provide a Data Management Strategy that addresses challenges and opportunities within the Agency

> › Objective 1 – Provide strategic direction to address recurring themes, business pain points and potential opportunities within the Agency related to data management

> › Objective 2 – Recommend frameworks, processes, technologies, and tools that provide a future vision for resolving recurring data management challenges

1.5 REFERENCED DOCUMENTS

This document incorporates input from review and access to multiple documents including:

- iTrace (the Third Party Administrator's current systems documentation tool)
- Agency IT SharePoint
- AHCA Division of Information Technology Strategic Plan (Fiscal Years 2017-20)
- North Highland Data Management Framework
- Strategic Enterprise Advisory Services (SEAS) S-3 Enterprise Systems Strategic Plan
- MITA Part II Chapter 2 Data Management Strategy

1.6 AHCA DIVISION OF INFORMATION TECHNOLOGY GUIDING PRINCIPLES

In 2013, the Agency Management forged eight guiding principles, listed below, to consider in all development or enhancement of systems. These Guiding Principles are reiterated in the Agency strategic plan for fiscal years 2017-2020 and must be considered in the design when creating a new system or significantly enhancing a legacy system.

1. Do not duplicate data – Whenever possible, data should be fed from a primary source and changed/edited at the source. For example, profile information needed for multiple systems should be maintained in one.

2. Go Paperless where possible – Paper may never be entirely eliminated. However, when it comes into this agency, we should make efforts to work the document electronically where possible and practical. Any item that is routinely passed through the Agency is a good candidate such as applications, forms and anything requiring signature.

3. Ensure accountability & security – Whenever practical, all systems should ensure accountability throughout. This means that the system is engineered such that any changes or queries can be logged, and that the users should be restricted to a least access necessary role. The Agency uses Windows authentication and a centralized Active Directory for users within our network.

4. Measure effectiveness – With each system designed to produce a product or output, the program area should have established performance measures designed to measure the effectiveness of the effort. For example, if a process from point A to point B should be completed in less than 30 days, then the system should be

designed to dashboard and/or produce reports to measure the percentage of items completed that fall within the standard.

5. Communicate with appropriate systems – All systems should be designed to communicate with other systems when appropriate. Data should be shared in a manner that is useful to the programs in meeting their missions.

6. Don't recreate the wheel – Be sure to consider Software as a Service (SaaS) solutions that can manage all data and infrastructure in a secure and backed up environment. This also helps to ensure that the total cost of ownership (TCO) has been considered in the options.

7. Consider Mobile Devices – Our long term strategic view should consider how mobile technology will impact our work and how we interact with our systems.

8. Customer Impact – Always consider the impacts to the customer in your system design and processes.

1.7 STRATEGIC TOPIC INVENTORY

This document provides guidance on many data management strategy topics. In the development of this deliverable, the SEAS Vendor created a Strategic Topic Inventory tool used to develop and communicate the Agency's direction on a wide range of data management strategy topics. The tool organizes topics into hierarchical taxonomy based logical groupings in areas of interest to strategic, programmatic, technology, and program management domains.

The Strategic Topic Inventory has many features to present and communicate a spectrum of strategic direction options considered across the spectrum of time for a specific topic. A summary chart can dynamically display the strategic direction for a specific topic across the time spectrum from current state direction to direction for future years. The Strategic Topic Inventory includes a field documenting a summary analysis that describes the context and considerations that influenced the defined strategy for each specific topic.

Extracts of the topic specific summary chart from the Strategic Topic Inventory tool are included throughout this document to communicate strategy and direction for many of the data management strategy decisions that are important for MES Program stakeholders to understand.

Over the course of the MES Program, the SEAS Vendor will continue to define and elaborate strategic direction on many data management strategy topics. The SEAS Vendor intends to continue to use the Strategic Topic Inventory tool as a discussion, recommendation, and communication vehicle for defining data management strategy direction as topics arise.

Below is a hyperlink to the active Strategic Topic Inventory on the MES Project Repository. All topics are available on the Topics sheet. Individual topics can be more easily viewed on the Chart sheet by entering the topic specific Displaying Row.

- SEAS Strategic Topic Inventory

The SEAS Vendor developed and maintains this Microsoft Excel based tool that resides as a document in the MES Projects Repository.

Exhibit 1-1: Strategic Topic Inventory Item Sample shows a screen shot example of a populated strategic topic.

Area:	Service Delivery - Technology		Description:				
Category:	Technology Strategy		What is the Agency tolerance for technology change and novelty?				
Sub-Category	Maturity						
Topic:	MES Technology Adoption						
Importance:	Medium	Strategy Status:	Draft Reviewed				
Displaying Row:	188						
Strategic Direction		Current	2018	2020	2022	2025	
Innovators							
Early Adopter			X	->			
Early Majority		AHCA IT / Non-FMMIS	->			X	
Late Majority		X					
Laggard							
Analysis:		The Rogers Bell curve categorized an organizations tolerance for change and novelty. The MES Technology adoption toward innovation will be tempered by a determination of the potential Agency disruption. They estimate organization tolerance for change and novelty across all industries as: Innovators (2.5%) Early Adopters (13.5%) Early Majority (34%) Late Majority (34%) Laggards (16%)					

Exhibit 1-1: Strategic Topic Inventory Item Sample

SECTION 2   ROLES AND RESPONSIBILITIES

This section identifies the roles and responsibilities for the primary stakeholders that maintain or use this document.

ROLE	RESPONSIBILITY
SEAS Vendor Data Architect	- Identifies the data management related technologies and processes necessary to improve the Medicaid Enterprise. - Propose data management solutions that align to MITA 3.0, State, and Agency specific Medicaid requirements. - Reviews and proposes new emerging data management technologies to the Agency. - Maintains the Agency Data Management Strategy. - Supports vendor procurements by providing information, extracts and details related to the Data Management Strategy.
Agency MES Technical Domain Lead	- Coordinates the participation of Agency stakeholders that identify data management strategy topics needing definition, recommendation or elaboration, review and provide feedback on proposed data management strategy topics. - Communicates data management strategy to Agency MES Domain Leads. - Supports MES Program leadership communication to Agency executive leadership. - Approves communications between the SEAS Vendor and MES Stakeholder Organizations related to MES Data Management Strategy. - Coordinates data management governance processes which are a mechanism for arbitrating differences.
MES Project Vendors (SEAS, ISIP, EDW, Module)	- Follows the strategic direction in the Data Management Strategy in proposing, discussing, and implementing technology for the Medicaid Enterprise. - When necessary, recommends data management technologies and solutions applicable to the implementation of MES Projects that align to MITA 3.0 and the Data Management Strategy.
MES Stakeholder Organizations	Reviews and as appropriate may align technology solutions with MES data standards, systems and processes per the Data Management Strategy to improve Medicaid program outcomes.

Exhibit 2-1: Roles and Responsibilities

SECTION 3 DATA MANAGEMENT AND INTEGRITY APPROACH

3.1 DATA MANAGEMENT VISION SUMMARY

Quality data, tools, and systems optimized for the Agency workforce is a foundational enabler to improve healthcare for all Floridians. This section defines the MES Data Management vision that aligns with the overall MES strategic priorities and aligns with the 2017-20 AHCA IT Strategic Plan. This vision guides the data management strategy allowing the Agency, MES Project Vendors, health plans, providers, and MES Stakeholder Organizations to improve recipient and provider experience in pursuing wellness for all Floridians.

While the Agency compares favorably to other states in low cost operations, recipient wellness and some recipient experience measures, the Agency has the opportunity to improve its organizational capability and potential. The people of the Agency demonstrate personal responsibility, ownership, and accountability to do their best to support providers and recipients. For many reasons this culture of accountability and ownership has led many business units to:

- Gather and store data
- Perform data edits, cleansing and transformation
- Profile and analyze data
- Produce reports
- Control and protect the security of the data

For many units, working independently with data is a necessity and not a preference. The current data and technology assets, culture, and processes of the Agency optimize processing from a discrete perspective compared to optimizing by leveraging the overall assets of the enterprise. The current data management culture constrains improved maturity and realization of organization potential. It also creates opportunities to improve the:

- Consistency of information and analysis provided in response to inquiries to the Agency
- Quality of analysis and depth of insights provided to the legislature, health plans, providers, and the public
- Protection of data and privacy
- Transparency and ease of access to program information
- Reuse of data and data sharing between agencies, systems, and business units

As the MES realizes its vision, data governance will evolve to confirm users have appropriate access to consistent, high quality data from a common source. The change to migrate from

information silos is a current necessity that is becoming increasingly critical to effect significant program cost and data protection improvements.

The MES Data Management Vision emphasizes six primary strategies that align with the overall MES strategic priorities:

- Improve data quality by operating from a single source of policy truth
- Evolve core processing with data validation at the point of business event data collection
- Provide seamless access to a real-time, 360-degree (360˚) view of recipient and provider information
- Decouple data from proprietary systems and application stores
- Operate with business area and persona optimized data marts and data analysis tools
- Prepare to collect and manage recipient and provider experience and outcome data

Improve data quality by operating from a single source of policy truth. Today, data edits, data validations, and data transformations are the electronic implementation of policy. The inconsistent application of data edits, validations, and transformations to the many different Agency data stores means there is no single source of policy truth which causes confusion and lack of trust in the data both within the Agency and with external consumers of Agency data. For example, data edit rules and policies are applied differently in the front-end of FMMIS Interchange when compared to the back-end resulting in claims rejections. Different business units and individuals implement policy by applying specific data edits, validations, and transformations to their own data sets to meet their needs or preferences. Often, separate systems support different versions of data validation and transformation. When each business area can claim common data is not right for the unit, this leads to many propagations of duplicated data and no true single source of the truth. The Agency's strategy is to centralize and standardize data edits, data validations, and data transformations applying the policy to a single source of truth data set. After consolidation, a single set of policies operationalized as system edits, validations, and transformations decreases the need for business unit or individual specific clones of data. After a single source of policy truth exists, health plans and providers can use the electronic implementation of this policy to validate information before submission to Agency systems reducing errors and rejects.

Evolve core processing with data validation at the point of business event data collection. Today, high-volume claims and encounter processing occurs in a single system that validates submissions in a complex, difficult to maintain claims processing engine. The current system is a stable, reliable workhorse that is essential for timely and accurate payments to health providers in Florida. Naturally, there is reluctance to introduce risk to this critical processing engine because of the transaction volumes and State spending processed by the system. However, evolution of core claims and encounter processing is essential for the Agency to meet its mission and strategic priorities. The most significant improvements in

provider experience, recipient experience, levels of fraud, and provider administrative costs depend on how core processing works. The Agency strategy is to evolve core processing by allowing health plans and providers to validate and verify claim and encounter data before submission to the Agency. Evolutions in core processing will reduce errors, rejected transactions, denied claims, and encounters and support costs. The Agency strategy to evolve core processing involves:

- Providing access to an electronic set of policy truth (e.g. implemented via rules engine)

- Providing health plans and providers with recipient, provider and reference data needed for evaluation against the electronic set of policy truth

- Having health plans and providers validate and resolve errors before claim and encounter submission by validating data at the point of business event. This will be accomplished through services the Agency will expose to health plans and providers allowing them to validate data against edit rules and policies prior to submitting to the Agency.

- Submitting validated claims and encounter records that can be accepted with minimal Agency processing

The Agency strategy of going beyond the boundaries of the Agency to fix data quality problems is foundational to address symptomatic and derivative issues that affect many business functions.

Provide seamless access to a real-time, 360˚ view of recipient and provider information.
Today, batch files drive most of Medicaid system processing. The Agency strategy is to use technology to assemble information in near real-time from all relevant sources to make processing decisions. The near real-time, 360˚ view of recipient information will eventually include information from other Medicaid stakeholder organizations providing access to comprehensive social determinants of care data. Access to current and complete recipient information will improve service authorization decisions, treatment, and enhance coordination of care by health plans and providers. The information will also help organizations in the community of care to deliver non-Medicaid services to recipients. Providers of education, child welfare, elder care, employment, and other services can be more effective by leveraging information and collaborating with other providers to benefit the recipient. For example, another state's analysis of behavioral issues in schools leading to class disruption, detention, suspension, and expensive behavioral services found the root cause often originates from health issues related to vision, hearing, and dental screenings. By sharing claim or encounter information with appropriate data privacy protections, educators may accommodate children, provide proper referrals, and confirm screenings occur. Providing service providers with real-time access to a comprehensive view of recipient information should also help the Agency, health plans, and providers to identify if increased coordination of care is prudent and justified.

Decouple data from proprietary systems and application stores. Today, FMMIS and most application systems use tightly coupled databases that contain information structured for use in

an individual application. The Agency data management strategy is to manage data as a service. New MES modules operate using data access services that connect to an operational data store that is independent of specific systems or modules. The operational data store provides data to applications through service calls or application programming interfaces (APIs) by subject areas, which is a commonly used and supported technical pattern. Decoupling data from proprietary systems and databases helps operate from a single source of the truth and reduces data duplication. This strategy simplifies access, improves security, and enables business agility to replace or improve a new module. Decoupling will also simplify the future migration to emerging virtual data access technologies (e.g. blockchain) that allow entire industry ecosystems to contribute data, access data and operate from a single secure information source.

Operate with business area and persona type optimized data marts and data analysis tools. As it relates to data strategy, a persona categorizes and defines the data and analytic usage and processing characteristics for a person. The persona generalizes the types and breadth of data used and processed and the types of tools used to perform a role. In most organizations, there are 5-10 different personas. Currently, several hundred Agency personnel routinely develop and execute custom Structured Query Language (SQL) queries in roles as power users. Moving forward, the Agency data strategy is to provide optimized data marts and tools that meet the needs of each combination of business area and data processing persona type. For example, users that perform advanced data scientist level analytics may need access to pull the data into more sophisticated software programs such as SAS to analyze the data more effectively. A data mart to support some personas would allow for a large download in a quick and efficient manner directly by the users themselves. This new strategy should reduce costs and improve responsiveness to business needs by rightsizing technology spend based on business persona need. See section 3.9 Business Intelligence (BI) for persona definitions.

Prepare to collect and manage recipient and provider experience and outcome data. Today, the Agency and entire healthcare industry has limited visibility to comprehensive recipient and provider experience or health outcome data. Survey and sampling provide limited feedback mainly about recipient satisfaction with provider interactions. Across all industries, system and process improvements are raising expectations of recipients and providers. The Agency expects increased scrutiny on the overall costs, time spent, and quality of service interaction by recipients and providers in the delivery of healthcare services. For the Agency, health plans, and providers this means collecting, storing, and analyzing more data and new types of data with new dimensions of analysis. Collecting experience data efficiently also requires new applications and technology. Likewise, emerging advanced payment models (e.g. Diagnosis Related Grouping (DRG), Enhanced Ambulatory Patient Grouping (EAPG), bundled payments) introduce changes to core claims and encounter processing systems.

3.1.1 DATA MANAGEMENT STRATEGY VISION TO-BE DIAGRAM

Supporting the MES Data Management Vision six primary strategies mentioned above, is the Data Management Vision To-Be diagram shown in **Exhibit 3-1: Data Management Strategy**

Vision To-Be Diagram and the Data Management Strategy Vision Enabling Capabilities shown in **Exhibit 3-2: DMS Vision Enabling Capability Descriptions**.

Exhibit 3-1: Data Management Strategy Vision To-Be Diagram provides a conceptual overview of major data management strategy vision enablers of the MES.
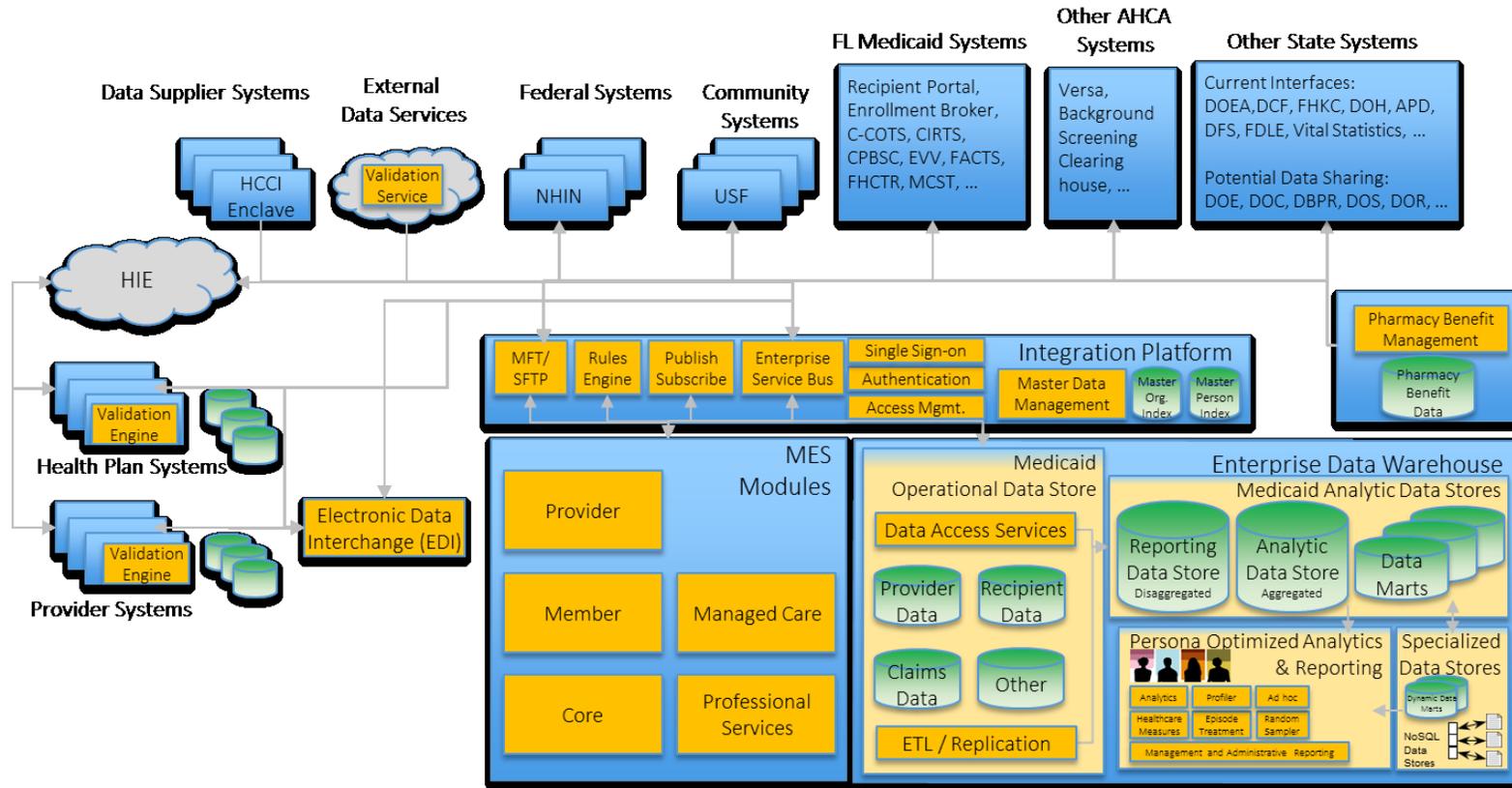
Exhibit 3-1: Data Management Strategy Vision To-Be Diagram

3.1.2 DATA MANAGEMENT STRATEGY VISION ENABLING CAPABILITIES

Exhibit 3-2: DMS Vision Enabling Capability Descriptions provides a brief description of each data management strategy enabling capability depicted on **Exhibit 3-1: Data Management Strategy Vision To-Be Diagram.**

ENABLING CAPABILITY	DESCRIPTION
Managed File Transfer (MFT)	Enables fast and secure transmission of files between systems
Rules Engine	Provides decisions based on edit rules, policy and datasets
Validation Service	Public or third-party service that validates pre-authorizations, claims and encounter transactions
Validation Engine	Processing engine within distributed plan and provider systems that validates and makes pre-authorization, claims and encounter acceptance decisions using rules and policy distributed by the Agency
Publish Subscribe	Notifies subscribers/designated systems of information updates about a recipient or provider
Enterprise Service Bus	Connects any approved request for data or processing to the data or processing service provider in real or near real time. Real time processing is continuous and typically happens in seconds. Near real time processing may not be continuous and typically happens in minutes rather than seconds. In addition, real time processing is synchronous which simplifies the request response process. Near real time processing implies asynchronous processing which adds the complexity of input queuing and accepting asynchronous responses.
Service Registry / Repository	Tracks web services and usage information
Single Sign-on	Allows users to authenticate to multiple systems using the same user id across systems
Authentication	Common framework that authenticates user access with modules and applications
Access Management	Common framework that manages role based access control within modules and applications
Master Person Index	Processing that identifies records about the same person within a system or found in other systems. Recommend using Commercial-off-the-shelf (COTS) Master Data Management (MDM) software to identify identity linkages.
Master Organization Index	Processing that identifies records about the same organization within a system or found in other systems. Recommend using COTS MDM software to identify identity linkages.
Master Data Management	System or rules that evaluates conflicting data about a person or organization to present a best or "golden record" which improves data quality and encourages data sharing through data content clarity.

ENABLING CAPABILITY	DESCRIPTION
Data Access Services	Provides decoupled access to data at varying levels of granularity. Data access services will span from elemental data services to module specific data services to composite cross module data services.
Operational Data Store	The data store of transactional data. Access to operational data is through data access services and APIs
Extract Transform Load (ETL) / Data Replication	Software that transfers information between data stores
Reporting Data Store	A data store optimized for use by dashboards and reporting and is continuously updated with data from the operational data store
Analytic Data Store	The data store optimized for analytic analysis. Also referred to as the data warehouse
Data Marts	Specialized data stores that are structured and optimized for specific types of analysis or used by specific business units
Dynamic Data Marts	Data stores that are created upon request in an optimized structure for a specific analysis or type of analysis
Not Only SQL (NoSQL) Analytic Data Stores	Analytic data store that is optimized for unstructured data sources and big data analytics

Exhibit 3-2: DMS Vision Enabling Capability Descriptions

Exhibit 3-3: DMS Vision Enabling Capabilities – Strategy Mapping shows each data management strategy mapped to the pillars of the Data Management Strategy Vision.

ENABLING CAPABILITY	SINGLE SOURCE OF TRUTH POLICY AND EDIT RULES TRUTH	DATA VALIDATE AT POINT OF DATA COLLECTION	SECURE REAL-TIME 360° VIEW OF INFORMATION	DECOUPLE DATA FROM APPLICATIONS	BUSINESS AND PERSONA OPTIMIZED DATA / TOOLS	RECIPIENT AND PROVIDER EXPERIENCE / OUTCOME DATA
Managed File Transfer (MFT)			✓			
Rules Engine	✓	✓				
Validation Engine	✓	✓				
Validation Service	✓	✓				

ENABLING CAPABILITY	SINGLE SOURCE OF TRUTH POLICY AND EDIT RULES TRUTH	DATA VALIDATE AT POINT OF DATA COLLECTION	SECURE REAL-TIME 360° VIEW OF INFORMATION	DECOUPLE DATA FROM APPLICATIONS	BUSINESS AND PERSONA OPTIMIZED DATA / TOOLS	RECIPIENT AND PROVIDER EXPERIENCE / OUTCOME DATA
Publish Subscribe			✓			
Enterprise Service Bus			✓			
Service Registry / Repository			✓			
Single Sign-on			✓			
Authentication			✓			
Access Management			✓			
Master Person Index			✓			
Master Organization Index			✓			
Master Data Management			✓			
Data Access Services		✓	✓	✓		
Operational Data Store				✓		
ETL / Data Replication				✓	✓	
Reporting Data Store				✓	✓	
Analytic Data Store					✓	✓
Data Marts					✓	✓
Dynamic Data Marts					✓	✓
NoSQL Analytic Data Stores					✓	✓

Exhibit 3-3: DMS Vision Enabling Capabilities – Strategy Mapping

The rest of this document provides strategic guidance and elaboration on topics organized around an enterprise data management framework. The data management framework provides structure and allows reuse of best practices for organizations of similar size, industry, and profile.

The MES Technology domain also has a documented MES Technical Management Strategy (SEAS Vendor Deliverable T-4: Technical Management Strategy) that supports the primary strategies of Data Management Strategy and provides strategic guidance and elaboration on topics related to each area in the data management framework. Together the MES Data Management Strategy and MES Technology Management Strategy communicate the strategic direction that will allow the Medicaid Enterprise including the Agency, MES Stakeholder organizations, MES Project Vendors, and the SEAS Vendor to help the Agency operate at optimal levels of efficiency and effectiveness achieving the Agency's potential.

3.2 DATA MANAGEMENT STRATEGY

The Data Management Strategy (DMS) provides guidance for future data systems and modernization of current enterprise data management systems. The strategy includes modernizing the Agency's data infrastructure to support the transformation of Agency business and application systems. Over the course of the MES Program, the Data Management Strategy will incrementally evolve to refine and provide additional guidance on data management strategic topics that benefit the MES Program.

The Data Management Strategy describes an approach to the overall management of the availability, usability, integrity, and security of the Agency data assets. The overall purpose of the Agency's Data Management Strategy is to:

- Make data integration efforts within and across agencies more efficient
- Support MITA's guidance for modularized implementation of various healthcare components and easier sharing of data
- Provide a common set of processes, tools, and data standards for the Agency's data solutions
- Improve data quality, reduce duplication, and associated overhead
- Comply with state and federal requirements
- Reduce technology support and maintenance cost
- Manage structured and unstructured, operational, transactional, reporting, and analytic data across the Agency

The first information systems to leverage the modernized enterprise data platform and processes for the Agency will be the Integration Services and Integration Platform (ISIP) that includes an Enterprise Service Bus (ESB) and the Enterprise Data Warehouse (EDW). These

two platforms provide the foundation for transforming the Agency into a data-driven organization and improving data quality, performance, and information accessibility.

Input from multiple business areas informed the overall Data Management Strategy. Agency personnel knowledgeable in integration with Medicaid Accounts Receivable (MAR), Agency contracted enrollment broker, Division of Operations, Bureau of Financial Services, claims payment funding and disbursements, recoupments, and compiling and publishing required federal reporting provided insights relevant to this initial definition of the DMS. Future iterations of the DMS will elaborate and refine the strategy as changes occur or more detailed direction is required.

3.3   MES Data Types

This section defines a framework of the types of data managed in the MES Program and prioritizes definition of management and integrity approaches for planned MES Projects. A MES Data Type is a category of data that has a unique profile or data management characteristics.

The MES Data Management Strategy generally applies to all MES Data Types. The uniqueness of some MES Data Types justifies additional analysis and elaboration of strategic direction, policy and techniques applicable to a specific MES Data Type. For example, contact center data recordings are an MES Data Type that has unique data management characteristics and opportunities for outcome benefiting usage (e.g. automated transcription of voice to text, analysis of speaker voice characteristics such as nervousness, or stress, use of voice patterns for authentication). For each specialized MES Data Type, future iterations of the MES DMS will provide additional strategy at the relevant point in time for use in MES Projects.

3.3.1   Inventory of MES Data Types

Exhibit 3-4: MES Data Types lists anticipated MES Data Types used and managed during the MES Program. The inventory provides a source for guidance and reference about MES Data Type specific considerations. Over time, as use of specific MES Data Type guidance is needed, the SEAS Vendor will provide strategic direction to enable stakeholders and MES Vendors to use and manage each type of MES data. The Characteristics Definition Priority describes the anticipated MES Project that would require definition of the strategic direction, policy, and techniques for each MES Data Type. Each Characteristics Definition Priority indicates <u>when</u> the corresponding data types will require definition of strategic direction, policy and techniques. Data Types identified as To Be Determined will be addressed in future iterations of the Data Management Strategy.

The types of processing performed in the MES Project categories below are:

- ▪ Core Processing – projects that primarily include claim and encounter processing and other core business processes

- Professional Services – projects that provide business and technical services that support other types of projects including call center, document management, communications and other reusable processes

- Integration Services and Integration Platform – projects that enable connectivity, data sharing, identity consolidation and security to enable integration and interoperability

- Enterprise Data Warehouse – projects related to data management, storage and analysis of enterprise data

DATA TYPE	CHARACTERISTICS DEFINITION PRIORITY
Claims Data	Core Processing Project(s)
Encounter Data	Core Processing Project(s)
Unstructured Operational Data	Core Processing Project(s)
Financial Data	Core Processing Projects(s)
Outbound Letters	Professional Services Project(s)
System User Authorization Data	Integration Services and Integration Platform Project
System User Access Control Data	Integration Services and Integration Platform Project
Person Identity Matching Data	Integration Services and Integration Platform Project
Organization Identity Data	Integration Services and Integration Platform Project
Reporting Data	Enterprise Data Warehouse Project
Analytic Data	Enterprise Data Warehouse Project
Electronic Health Records	To Be Determined
Clinical Health Data (e.g. Genetic, health sensor, image, video, etc.)	To Be Determined
Federal Grantor	To Be Determined
Court Records	To Be Determined
Investigation Data	To Be Determined

DATA TYPE	CHARACTERISTICS DEFINITION PRIORITY
Customer Communication Data (e.g. Short Message Service (SMS) Outbound, Contact Center Recordings, Chat Session, Social Media, Agency Email etc.)	Professional Services Project(s)
Business Rules and Policy Data	Integration Services and Integration Platform Project
Statute and Rule Data	Professional Services Project(s)
Web Published Static Content	Professional Services Project(s)
Natural Language Data	Professional Services Project(s)
Language Translations	Professional Services Project(s)
Geographic Information System (GIS) Data	Enterprise Data Warehouse Project

Exhibit 3-4: MES Data Types

3.3.2 MES Data Type Characteristics

Exhibit 3-5: MES Data Types Characteristics is a preliminary list of the types of information that will describe each MES Data Type. The specific data type characteristic field types, format, and validations will be specified when the MES Data Type List is defined in the MES Repository.

DATA TYPE CHARACTERISTIC	DESCRIPTION
Data Type Name	Short Name of the data type
Description	Description of the primary characteristics of the data type
Usage	Highlights of usage profile of the data type including primary users, types and numbers of systems that use this data type
Volume	Descriptive summary of approximate volume for this data type
Status	Status of the definition of characteristics of the data type
Severity	Relative importance of this data type
Data Type Category	Categorization that groups related data types (e.g. operational, contact center)
Data Modeling Strategy	Strategy to model this type of information and definition of unique data modeling considerations
Storage Strategy	Strategy used to store this data type

Data Type Characteristic	Description
Archive Strategy	Tracks web services and usage information
Data Classification	Data security classification and unique considerations for this data type
Data Owner	Description of data ownership for this data type
Technology Strategy	Description of technology strategy and tools used to manage this data type
Access Control Strategy	Data authentication and access control considerations for this data type
Usage Accounting	Description about usage tracking, cost accounting considerations
Monetization	Description of monetization policy and restrictions for this data type
Data Type Owner	The person or role that is responsible for MES Program definition and use of this data type
Relevant Standards	Linkage to MES Standards applicable to management and use of this type of MES data

Exhibit 3-5: MES Data Types Characteristics

3.3.3 MES Data Types Strategy Communication and Maintenance

Stakeholders to the MES Program including the Agency, SEAS Vendor, IV&V Vendor, MES Stakeholder Organizations, and MES Project Vendors will reference the Data Type descriptions and strategies throughout the MES Program.

The strategy to communicate and maintain this information is to create a custom MES Data Type Strategy list on the MES Repository. The MES Data Type Strategy list will become the system of record for the data type characteristics for each defined MES Data Type. The MES Data Type Strategy list will include status fields that allow users to understand the maturity of each data type strategy record.

3.4 DATA MANAGEMENT CAPABILITY FRAMEWORK
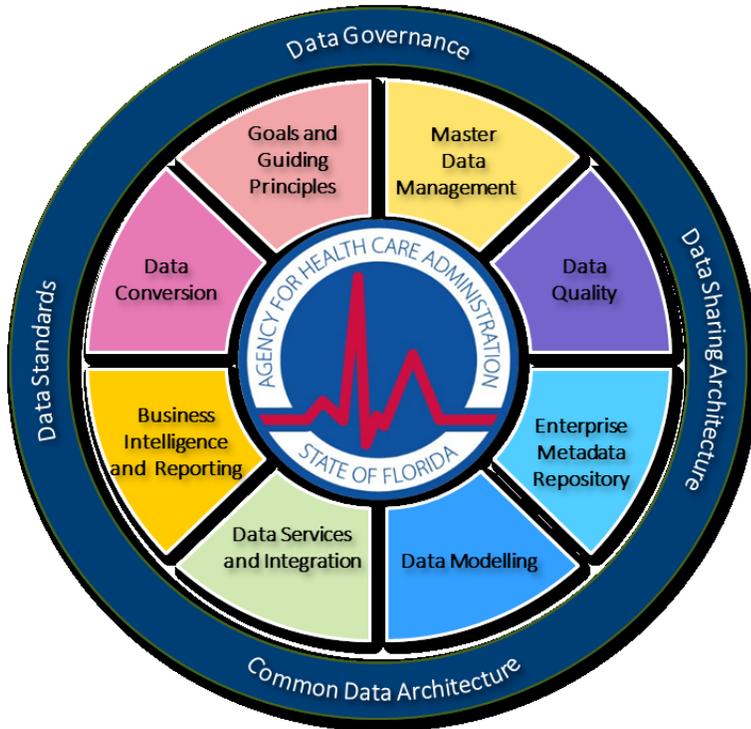
Exhibit 3-6: Data Management Capability Framework depicts the capabilities of the overall capability framework.
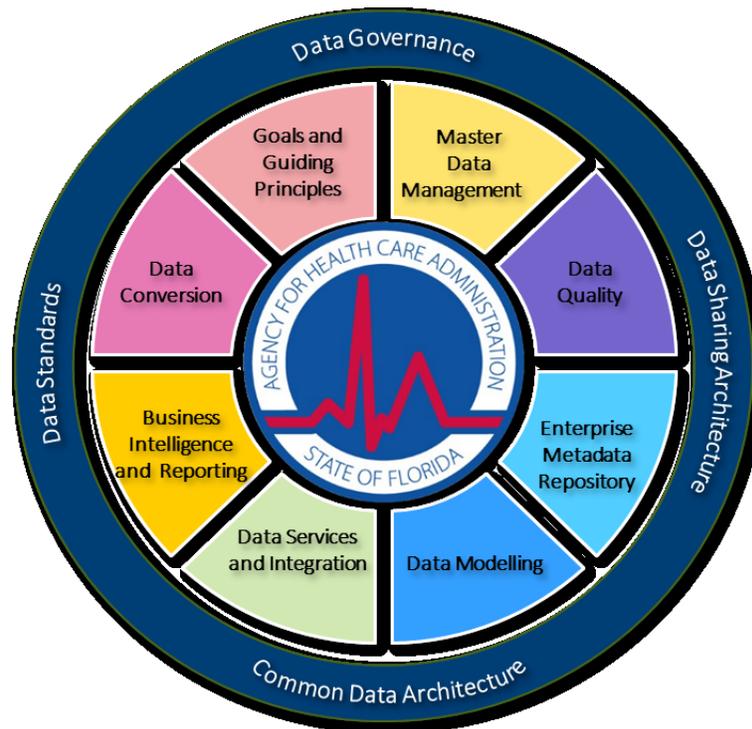
Exhibit 3-6: Data Management Capability Framework

The following bullets further describe the capabilities outlined in **Exhibit 3-6: Data Management Capability Framework:**

- **Data Governance:** Aligns the organization with the business strategy and enables prioritization and approval of initiatives and a plan for the introduction of processes that will continuously monitor and improve data quality.

- **Master Data Management:** The implementation of repeatable sets of business rules and the supporting data management and data distribution systems that define the value, content, and structure of specific data and data attributes. These data require consistent use by all systems within a given scope of the Agency.

- **Data Quality:** The processes for verifying data within source systems and following standards so that business rules are in place to govern the use and movement of data.

- **Data Sharing Architecture:** The framework that enables sharing of the underlying data with MES stakeholders through reusable data services.

- **Enterprise Metadata Repository:** A taxonomy for the effective definition and organization of the Agency's metadata, which provides valuable ways to view and report on information that drives decisions and actions.

- **Common Data Architecture:** An efficient process and infrastructure for common data repositories and services.

- **Business Intelligence and Reporting:** The process of converting transaction or production
information into useful knowledge via available reporting tools for real-time (dashboard), snapshot (scorecard), and detailed data display (reporting).

- **Data Standards:** A set of rules that must be followed to make decisions about data and data management solutions. Data standards guide the implementation of each of the modules, key data domains, and data management capabilities within the Agency.

- **Data Services and Integration:** The process of "extracting" data from internal and external sources, "transforming" it to fit business needs, and "loading" it into appropriate systems, often utilizing an ETL (Extract, Transform, Load) tool.

- **Data Conversion:** The process and systems to transform data from source systems and formats to target systems and formats. Data conversion is both a one-time activity as part of system migration as well as an ongoing activity related to the ingestion (input) and publishing (output) of data.

- **Data Modeling:** The process of creating, maintaining, and sharing data models necessary to support the information systems within an organization.

- **Goals and Guiding Principles:** The development of strategic direction and implementation planning for the acquisition, storage, and use of data. This includes all data types.

3.5 DATA MANAGEMENT GOALS

The Agency's Data Management Strategy includes the following Goals and Objectives.

- Goal 1: Provide complete, accurate and consistent healthcare information across all internal divisions, processes, and MES stakeholders.
    › Objective 1: Enable more timely responses to meet regulatory demands (CMS, MITA, etc.).
    › Objective 2: Enable a 360° view of key master data information such as provider and recipient.
    › Objective 3: Enable flexibility and easier sharing of healthcare information through well-defined enterprise modeling and metadata management processes and guidelines.

- Goal 2: Promote reuse of existing and new services that enable interoperability and modularity as key principles and provide cost effective high performing data services and solutions through a well-designed common data architecture and data sharing architecture.

- › Objective 1: Promote reuse and maximize the effective and efficient use of data and information assets.
- › Objective 2: Enable both technology and business process consistency through reuse of common tools and processes through data services and integration hub.
- › Objective 3: Reduce data duplication and redundant data management processes and environments.
- Goal 3: Provide increased insight from data.
  - › Objective 1: Build a well-integrated healthcare infrastructure that provides positive healthcare information throughout its divisions and across agencies.
  - › Objective 2: Provide increased insight to key partners enabling better fraud analytics and decision-making.

3.6 DATA MANAGEMENT GUIDING PRINCIPLES

The Agency's Data Management Strategy applies the following guiding principles.

- **Decouple information and application processing** - Applications design and implementation uses highly decoupled, autonomous services enabling applications and systems to perform tasks independently. Design applications to access business data processing using APIs and design data access services using Object Relationship Modeling techniques to promote information decoupling and object-oriented principles.

- **Reduce creation of multiple common person identifiers** - The Master Person Index is the system of record to link person identity records in different systems. Avoid approaches to assign multiple redundant person identifiers used for cross system identity matching.

- **Use data replication appropriately** - Overuse of data replication can lead to increased support, higher maintenance costs, and data synchronization and governance challenges. Reduce data replication whenever possible. In the Agency's case, an example of appropriate data replication is to maintain the data integrity of the new ODS and legacy FMMIS data repositories throughout the modular transition.

- **Real-time Information access** - Data consumers and decision-making systems drive the need for real-time information access. While traditional approaches to data integration, such as ETL and data replication are needed as part of the overall data management strategy, systems and modules should be designed with processes that enable real-time access to information whenever possible.

- **Reusability** - The integration architecture and services must be designed with reusability in mind. The goal is to build tools, interfaces, and data services that can be leveraged in a standardized and consistent way across the Agency. This will help reduce operational cost and overhead.

- **Data-as-a-Service (DaaS) vendors** - To manage the volume of data in an increasingly connected world, systems and processes should consider leveraging DaaS vendors. Data as a Service is a business-centric service that transforms raw data into meaningful and reusable data assets and delivers these data assets on-demand via a standard connectivity protocol in a pre-determined, configurable format, and frequency for internal and external consumption. It is an emerging architectural pattern where an external vendor manages a topic or topics of data for the business and simply provides an interface, which could be Service-Oriented Architecture (SOA) based or REST (Representational State Transfer) API based.

- **Interoperable** - Interoperability is the ability for different systems and applications to communicate, exchange data, and use the information that has been exchanged. All modules and services should be designed and implemented with interoperability as a key architectural goal.

- **Manage Information as an asset** - With the need for and value of data in day-to-day business operations and strategic decision making, it is imperative that each data service is managed as an asset with distinct business value.

- **Data governance and data stewardship shared responsibility** - While data governance and data stewardship are distinctly different, their importance and connected relationship necessitate that both IT and the Business need to be an equal partner to be successful in implementing data governance processes. Enroll and involve both business and technology leaders to support any large transformation effort.

- **Apply key data quality processes** - Data accurately represents reality, so do not modify data without going through a data governance, remediation, and data resolution process.

- **Maintain data security** - Apply encryption and department approved security protocols and processes to all data managed within the Agency. Maintain data confidentiality to prevent disclosure to unauthorized persons or systems. SEAS deliverable T-8 Enterprise Data Security Plan contains additional information about MES Data Security.

- **Comply with data standards** - Data Standards enable better data governance and data management. Implement and enforce appropriate data standards as relevant to each module. Further details on data standards can be found in SEAS deliverable T-3 Data Standards.

Agency direction on strategic topics related to MES Guiding Principles follows:

Strategic Topic 3-1: Data as a Service (DaaS) Vendor Strategy describes the Agency strategy for leveraging DaaS vendors.

DAAS VENDOR STRATEGY	Current	2018	TIMELINE 2020	2022	2025
No DaaS vendors	X	X			

DaaS VENDOR STRATEGY	Current	2018	TIMELINE 2020	2022	2025
DaaS vendor for Non-Critical Data					
DaaS use case(s) implemented as part of modularized implementation			X		
Expand and align DaaS vendor use to overall data federation strategy				X	->
Analysis	- Data as a Service is a business-centric service that transforms raw data into meaningful and reusable data assets, and delivers these data assets on-demand via a standard connectivity protocol in a pre-determined, configurable format and frequency for internal and external consumption. It is an emerging architectural pattern where an external vendor manages a topic or topics of data for the business and simply provides an interface, which could be Service-Oriented Architecture (SOA) based or REST (Representational State Transfer) API based. - Currently, HQA is using the Health Care Cost Institute (HCCI) Data Enclave from NORC at the University of Chicago for the Florida health price finder. This data enclave model could be extended to a DaaS paradigm and used as a pattern for future engagements and data needs within the Agency. - DaaS is an established pattern frequently used in Customer Resource Management (CRM) and Enterprise Resource Planning (ERP) applications but as more and more vendors emerge in the healthcare industry offering DaaS solutions, this could become an attractive option for Agency consideration. The State of California's 2013 Department of Healthcare Services Strategic Vision also recommends DaaS as a component of their overall information architecture.				

Strategic Topic 3-1: Data as a Service (DaaS) Vendor Strategy

3.7   MASTER DATA MANAGEMENT APPROACH

Section 5 Common Data Architecture explains further the Master Data Management capabilities, tools, processes, and recommendations.

3.8   DATA QUALITY AND DATA INTEGRITY APPROACH

Data quality and integrity are especially important in the realm of healthcare information. As it relates to the Agency, data forms the basis for critical decisions like the payment of claims,

provider licensure, services provided to recipients, rate setting, and deep analytics. Inaccurate, late, or inaccessible data can quickly lead to costly and even damaging mistakes by the Agency, health plans, and providers.

3.8.1 DATA QUALITY FRAMEWORK

The Data Quality Framework provides structure to assess and measure data assets of the enterprise. **Exhibit 3-7: Data Quality Framework** shows four categories and characteristics of data quality.

Accuracy
- Believable
- Objective
- Complete

Relevancy
- Timely
- Value Add
- Quantity

Data Quality Framework

Accessibility
- Secure
- Access Methods

Representation
- Understandable
- Concise
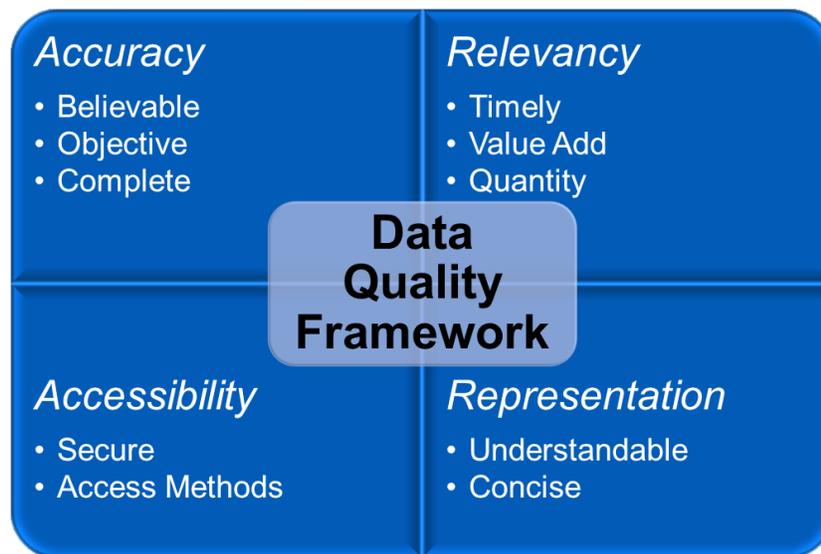
Exhibit 3-7: Data Quality Framework

This Data Quality Framework can be an input to the MITA State Self-Assessment evaluation of data quality within the Agency. The four categories of data quality are:

Accuracy

Data accuracy focuses on the correctness of data within the Enterprise. Is the data reliable and with the right level of precision for it to be considered accurate and usable? Data accuracy has the following three components:

- **Believable** – Is the data content credible and realistic?

- **Objective** – Is the data unbiased?

- **Complete** – Is the appropriate scope of information contained with the breadth and depth of the data?

<u>Relevancy</u>

Data relevancy is about the applicability and usability of data within the Agency. Data relevancy has the following three components:

- **Timely** – Is the data current enough to be usable and relevant?
- **Value Add** – Does the data provide operational value or a competitive edge?
- **Quantity** – Is there enough data available to meet the desired need and not too much data to slow down or impede processing?

<u>Accessibility</u>

Data accessibility is about the availability of the data and the ability to retrieve the data when needed. Data accessibility has the following two components:

- **Secure** – Is access to the data restricted accordingly to data classification and user role?
- **Methods of Access** – Is the data accessible through tools, systems and processes that support distinct business needs?

<u>Representation</u>

Data representation focuses on how easy it is to interpret the data. Data representation has the following two components:

- **Understandable** – Is the data logical and clear?
- **Concise** – Is the data well organized, compact and to the point?

The Data Quality Framework is a mechanism for the Agency to measure and assess data quality. This Framework can be an input to the MITA State Self-Assessment evaluation of data quality within the Agency and will be used to assess data quality during and after Modular implementations. These before and after data quality assessment scorecards will help assess the effectiveness of Modular implementations. The SEAS Vendor will work with the Agency to use the data quality framework to make data management decisions during the MES Project. This initial data quality framework implementation will use a simple scale (high, medium, low, or not applicable) to assess each category and characteristic in the data quality framework when assessing data quality.

The SEAS Vendor will aid in the ongoing evolution and adoption of the Data Quality Framework. The SEAS Vendor works with the MES Agency Technical Domain Lead and MES Technology Governance (MTG) of the MES governance structure to protect and build value from enterprise data assets using the Data Quality Framework.

Recognizing that the data quality framework adoption and maturity will happen over time, interim data quality measures should be taken through partial adoption of the framework and subsequent assessment. To build an initial data quality plan, the SEAS Vendor will work with the MES Agency Technical Domain Lead as well as the MTG to provide an initial Agency assessment against the ten data quality metrics contained in the four data quality categories. This initial assessment and subsequent evaluation will focus on identifying the metrics and categories where the biggest return on investment can be achieved in improving data quality on an interim basis until the full data quality framework is able to be adopted.

MES Program direction on strategic topics related to the Data Quality Framework includes:

Strategic Topic 3-2: Data Access Control Strategy describes the Agency strategy for controlling access to data.

DATA ACCESS CONTROL	Current	2018	TIMELINE 2020	2022	2025
Ad hoc - No Agency wide data access strategy. Individual systems control data access.	X	X			
Data access control is managed at the database level					
Data access control is managed at the data service level					
Data access control is managed at the application level					
Hybrid - Data access is controlled using a tiered approach			Role Based at Database / Data Service / Application	->	

DATA ACCESS CONTROL	Current	2018	TIMELINE 2020	2022	2025
Analysis	In the current MES environment, there is no data access strategy. Individual system owners control data access processes for individual systems. In the future state, it is expected that data access control is managed at the database, data service and application level using a centralized role based security paradigm. Single sign on, authentication, and access management are all enabled through the ISIP. Data access controls will implement defense in depth techniques that include system to system level control, user identity authentication to access application systems, user role based controls managed with applications. At the integration layer, policy based controls will control access to systems, services, service methods, and content. In all cases, logging of access requests will support identification of the requester including user id information.				

Strategic Topic 3-2: Data Access Control Strategy

3.8.2 DATA QUALITY MATCHING TECHNIQUES AND PROCESS

Until the MES enterprise achieves the goal of operating with data that provides a single source of truth, data quality variations will exist between silos of information. The implementation and use of the MES operational data store with data service access methods and the creation of industrialized, centralized or consistent policy, business rules, data edits, data validations, and data transformations will improve data quality and consistency.

While duplicate data exists in different data stores and where a single source of truth can be difficult to determine, the following data matching technique and process may help improve data quality:

▪ Identify the key data elements that data matching will be based on. This may be a unique identifier, an address, etc. or a combination of multiple unique identifiers (e.g. fully qualified name and address combined).

▪ Establish business rules to promote data quality, especially for the data elements identified as data matching criteria. Typical business rules will be data validation in terms of ranges, masking, type, data duplication constraints, and preemptive rules enforcement at the point of data collection.

▪ Identify data outliers that either violate the established data quality business rules or do not match when linked to peer data sets via the established key data elements. To address data outlier issues requires either removing the exceptions from the dataset, fixing the source data at the point of data collection, or leaving the data in the dataset but limiting their use or flag their use as outliers. Generally fixing data outliers at the point of data collection provides the most effective improvement in data quality.

3.9 BUSINESS INTELLIGENCE (BI)

Business Intelligence (BI) is a set of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information to enable more effective strategic, tactical, and operational insights and decision-making. The Agency's strategy is to use a Business Intelligence Framework (BI Framework) to improve the process of satisfying requests for reporting, dashboard, or analytical needs. The BI Framework organizes and aligns the strategy, people, processes, tools, and data to support the requirements and needs of Agency business units and individuals. **Exhibit 3-8: Business Intelligence Framework** depicts the major components of the BI Framework that will help strategically align projects to meet the business intelligence needs within the Agency. Strategically aligned projects produce focused actionable insights that help achieve the Agency's mission.
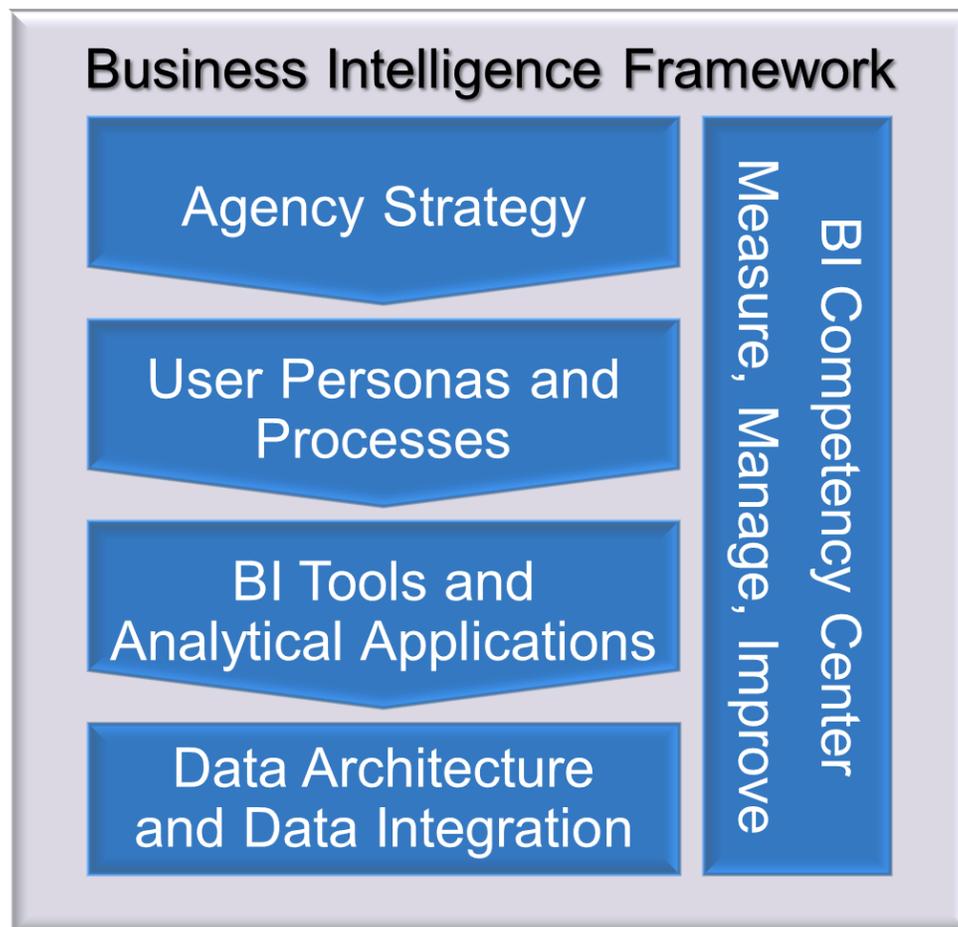
Exhibit 3-8: Business Intelligence Framework

The Business Intelligence Framework is a top-down model where each area informs the next area in the framework. The key components of the Business Intelligence Framework are:

Agency Strategy

Agency strategy and business needs serve as the overarching set of guiding principles for all Business Intelligence decisions. Agency strategies and objectives are defined, documented, and communicated within the Agency and used as the starting point of principles that inform decisions regarding BI processes, tools and underlying data architecture and data integration.

User Personas and Processes

Defining user personas and the corresponding processes used by each persona informs the BI tool selection process to confirm BI tools support the enterprise and the Data Architecture. Within the Agency, the existing types of user personas include:

- **Operational User** – Views predefined reports. Primarily seeks discrete numbers, pieces of information, and insights pertaining to their job function.

- **Data Retriever** – Collects information from one or more sources. Collected information provides others actionable information for processing. The collection process may gather information through several iterations of selecting data facts, filtering, transforming, and sub setting data before finding the insight they need.

- **Analysts** – Begins by selecting the data they need, typically from a larger dataset and then enriching the data by creating calculated fields and possibly combining the data with additional datasets. Often will create comparisons between data and evaluate trends across time.

- **Advanced Analyst** – Creates models for the Agency using advanced analytic tools like SAS, SPSS or R. Pulls data from multiple data sources both within the Agency and outside the Agency. Merges varied data sources together for analysis.

Each type of user personas has a unique set of processes based on their individual set of tools, data sources and deliverables for the Agency. The goal of the Business Intelligence Framework is to avoid standardizing on one set of processes across the entire enterprise but recognize the uniqueness within each persona and develop appropriate repeatable processes for each.

BI Tools and Analytical Applications

The MES strategy to select and evaluate Commercial-off-the-shelf (COTS) Business Intelligence and advanced analytic tools considers the Agency strategy and user personas. The BI tools strategy is to select a standard set of BI tools specific to business unit and user personas requirements. The Agency plans to select a suite of BI tools to improve satisfaction and effective usability across the user community. Standardizing on one size fits all tools would

result in significant dissatisfaction and inefficiency across the user community and lead to unnecessary tool fragmentation.

The Agency currently has multiple COTS BI and advanced analytic tools deployed including Tableau, SSRS, Power BI, BusinessObjects, SPSS, SAS, and R. Some tool rationalization across the Agency is warranted as Tableau, Power BI and BusinessObjects all provide similar capabilities and SPSS, SAS and R do so as well. Typically, organizations would have one, or possibly two, BI and analytic tools associated with each user persona. Limiting the number of BI and analytic tools available to each user persona type will help reduce acquisition and support costs and increase resource interchangeability (fungibility). Standardized tools allow for skill fungibility across departments, scale in license purchasing, improved training and an overall cost reduction. Confirmation of existing BI tools or selection of replacement or additional BI tool specific to will occur as part of the Enterprise Data Warehouse procurement and MES Module procurements.

Enforcement of BI COTS tool strategy is a shared responsibility across technology, purchasing/finance, and department level managers. Executive leadership plays an important role in coordinating tool decisions to achieve the benefits of standardization without constraining business units from maximum effectiveness. A Chief Data Officer (CDO) role can provide management and analytical expertise to implement the data tool strategy in the organization. For the MES Program, the SEAS Vendor will perform periodic reviews of BI COTS tool usage for each user persona type according to the Business Intelligence Framework.

Strategic Topic 3-3: Business Intelligence Tool Strategy describes the recommended business intelligence COTS tools strategy.

BUSINESS INTELLIGENCE TOOLS	Current	2018	TIMELINE 2020	2022	2025
Bring your own tool	X	X			
Multiple authorized tools			Exception Only	->	
Department specified tools by persona			Exception Only	->	
Agency specified tool suite by persona			X	->	
One size fits all					

BUSINESS INTELLIGENCE TOOLS	Current	2018	TIMELINE 2020	2022	2025
Analysis	Currently, the Agency has multiple authorized BI COTS tools (e.g. Tableau, SSRS, Power BI, BusinessObjects, SPSS, SAS, R, and Excel). The future state strategy is to standardize on one BI tool per persona. Standardized tools allow for skill fungibility across departments, scale in license purchasing, improved training and an overall cost reduction. The EDW/Data Analytics vendor will evaluate and propose a BI COTS tool suite to the Agency. Enforcement of BI COTS tool creep will be a shared responsibility across technology, purchasing/finance, and department level managers. Periodic reviews of BI COTS tools by persona will be conducted according to the Business Intelligence Framework as specified in the MES Data Management Strategy.				

Strategic Topic 3-3: Business Intelligence Tool Strategy

Data Architecture and Data Integration

The data architecture and data integration component of the BI Framework represent the foundation for all BI and advanced analytic activities and processes. This includes ensuring:

- The infrastructure is rightsized
- Appropriate datasets are available (via operational, reporting data stores, analytic data stores, and data marts)
- Data integration requirements are accounted for
- Performance considerations and requirements are defined

The SEAS Vendor collects and documents these non-functional requirements for each business area and user persona combination. The BI Framework will incorporate use of both the Integration Services and Integration Platform (ISIP) and future EDW capabilities. A combination of the Integration Services Platform and the EDW will be key components in integrating the multiple disparate data sets used within the Agency. The Analyst and Advanced Analyst user personas most often perform analysis of integrated information datasets.

BI Competency Center

A strategic direction is to establish a Business Intelligence Competency Center (BICC) that provides planning, measurement and support of business intelligence maturity improvement. The BICC is not a centralization of business intelligence users or standardization of all business intelligence processes. The BICC is responsible for developing the strategic plan and priorities for how the Agency supports BI. BICC are the owners and developers of the BI

Framework. The Agency currently has a BICC within the Health Quality Assurance (HQA) division. The BICC focus and impact is primarily limited to HQA. The BICC's scope should be expanded outside HQA to drive BI consistency in the areas of tools, processes, and shared data across the Agency.

Strategic Topic 3-4: Business Intelligence Strategy shows the recommended business intelligence strategy roadmap recognizing current state, future modular implementation, and enterprise data warehouse strategy.

BUSINESS INTELLIGENCE STRATEGY	Current	2018	TIMELINE 2020	2022	2025
Operational Data Store (ODS)	FMMIS / Non-Medicaid Stores	FMMIS / Non-Medicaid Stores	FMMIS / MES ODS	MES ODS	->
Reporting Data Store (RDS). ODS optimized for reporting (e.g. de-normalization)			X	->	
Centralized data warehouse	DSS / Non-Medicaid Warehouse(s)	DSS / Non-Medicaid Warehouse(s)	EDW	->	
Fixed data marts	X	X	Business function & persona optimized, Medicaid Accounts Receivable (MAR) reporting	->	
Department managed BI stores	X	X	Exception Only	->	
Dynamic data marts			Evaluate / Pilot	->	

BUSINESS INTELLIGENCE STRATEGY	Current	2018	TIMELINE 2020	2022	2025
Analysis	Currently, the Agency has department managed BI stores, some fixed data marts, and a centralized data warehouse. The future state strategy maintains unique analytic and reporting data stores optimized for business function and persona needs. These analytic and reporting data stores include a new operational Reporting Data Store (RDS), Enterprise Data Warehouse and Data Marts. Data Mart creation does not change data content, only data structure or format. This maintains consistency from operational data through reporting and analytic data. Department specific copies of data stores that have modified content from base operational data will be eliminated.				

Strategic Topic 3-4: Business Intelligence Strategy

3.10 METADATA MANAGEMENT

Section 7 – Enterprise Metadata Repository describes metadata capabilities, metadata repository management processes, and metadata tools.

3.11 DATA MODELING

Section 6 – Enterprise Modeling describes enterprise data modeling capabilities, data modeling governance, and tools.

3.12 DATA ARCHITECTURE AND ITS ROLE FOR DMS

Section 5 – Common Data Architecture and Section 8 – Data Sharing Architecture further describe the Agency's Common Data Architecture and Data Sharing Architecture.

3.13 DATA CONVERSION

Data conversion is the process of performing a one-time migration of data in one or more data stores or systems to a different system or data store. The migration process can include data use analysis, data cleansing, data format and value changes, and iterative mock conversion testing to prepare data for use in the to-be data stores. Data conversion is a significant transition activity in the implementation of the MES Program. The primary types of data conversion that will occur are:

- FMMIS databases to the MES Operational Data Store (ODS)
- Many standalone application data stores to use the ODS
- Multiple agency systems into consolidated data stores (e.g. licensing data)

- Historical DSS data no longer available in FMMIS to the Data Warehouse and Data Marts
- Consolidation of Content stores (e.g. document management data)

On an ongoing basis, data in some MES data stores will populate information in other MES data stores. The MES data store population processes extract, transform, and load information to other MES data stores that are optimized for use by specific systems or data use personas. Because MES data store population occurs as an ongoing activity, it is not technically considered data conversion even though MES data store population use many of the data conversion processes and quality controls performed in one time data conversion projects. Examples of ongoing MES data store population include:

- ODS data populated to the Reporting Data Store (RDS)
- RDS data populated to the Enterprise Data Warehouse
- Enterprise Data Warehouse data populated to Data Marts

This section highlights basic data conversion practices to follow during the build and implementation of MES modules. The SEAS Vendor will elaborate and industrialize data conversion practices to provide consistency and reduce transition risk related to MES Project implementations. While the data conversion from and to each system requires unique analysis and planning, the large number of data conversions justify standardization and process consistency.

3.13.1 DATA CONVERSION PRACTICE: CONVERTED DATA USAGE

Understanding the use of FMMIS data in the MES Module System is foundational to selecting the appropriate data conversion strategy, level of investment, level of quality, testing approach, tools and data architecture for the converted data. Realistic consideration of converted use allows the Agency to make informed decisions about the level of investment that is justified in the data conversion process. Characteristics of usage consider:

- How frequently the converted data is accessed
- What types of usage and user personas access converted data
- How many different people access the converted data
- How quickly converted data needs to be accessed

These and other analyses help select a data conversion strategy.

3.13.2 DATA CONVERSION PRACTICE: DATA QUALITY EXPECTATIONS

Managing converted data quality expectations is one of most important data conversion considerations and serves as a primary success factor in evaluating if the data conversion was successful. Managing converted data quality expectations involves accurately understanding the as-is data quality and articulating the to-be data quality.

Profiling existing data with automated profiling tools provides transparency to the quality of source data. This process helps identify data exceptions which could have implications on application and system logic.

The as-is data profile is an input to define and agree upon the to-be data quality for converted data. A data conversion project frequently brings a focus on data quality and often is a logical point to make data quality improvements. Often automated transformation can standardize data and correct inconsistencies in data values. Sometimes manual data cleanup occurs in source or target systems as part of the data conversion process. It is not always appropriate to invest in improving the quality of converted data. Likewise, the appropriate level of expenditures on testing and perfecting the data conversion process will depend on many factors including future usage, size of data, type of data and other factors. Formally defining data quality expectations provides transparency for better decision making and effective stakeholder communications.

3.13.3 DATA CONVERSION PRACTICE: MOCK CONVERSION AND TESTING

MES Project Vendors will perform pre-conversion testing to set a baseline and post-conversion testing to validate against the baseline. Mock conversions will be conducted to test the entire process of data conversion. The data conversion strategy will define an appropriate number of mock conversions and the level of testing with converted data.

Minimally, conversion error logs will be evaluated, record counts confirmed, and checksums validated. Multiple mock data conversion runs will help iterate through data conversion challenges and determine the necessary cutover period for the final go-live data conversion.

Following final go-live data conversion and testing, system users perform post migration User Acceptance Testing (UAT) as the final form of validation to confirm data quality.

3.13.4 DATA CONVERSION PRACTICE: AUTOMATED CONVERSION

The size and complexity of data assets at the Agency suggests automated data conversion is the default. While manual data conversions are acceptable for small, simplistic data sets, they are unacceptable for large, complex data sets. Developing a complex data conversion process is iterative, must be repeatable, and needs to execute in a timely manner. All these attributes necessitate automated data conversion processes for the Agency.

3.13.5 DATA CONVERSION PRACTICE: FLEXIBILITY

Because most Agency's data conversions use automated conversion tools and processes, the tools and processes required to accomplish the automation will require flexibility. The Agency's complex data landscape and modular approach to implementation will likely require the use of both COTS ETL tools and custom-built data conversion scripts to successfully complete the necessary data conversions.

3.13.6 DATA CONVERSION PRACTICE: LINK TO MASTER DATA MANAGEMENT

Data conversion and master data management considerations require evaluation when planning and executing a large-scale data conversion effort. Throughout the process of data conversion, master data will be identified which presents an opportunity for the Agency to profile, classify and begin maintaining master data policy. Section 3.7 Master Data Management Approach contains additional information on Master Data Management.

3.13.7 DATA CONVERSION PRINCIPLE: HISTORIC AND INQUIRY ONLY DATA

Historic and inquiry only data justify special consideration. Conversion of structured source data to unstructured data stores (e.g. NoSQL) can be a technique to migrate data without performing the rigorous field by field validation and transformation processes required for active transactional data. This technique does not apply to all structured source data and use should carefully consider how historic and inquiry only data will be used.

SECTION 4 ENTERPRISE DATA MANAGEMENT AND DATA STEWARDSHIP

4.1 DATA MANAGEMENT AND DATA STEWARDSHIP

Per MITA 3.0 Part 2 Chapter 2, "Data management and data stewardship implement Data Governance, Data Stewards, Data Owners, and Data Policy. Data Governance defines the governance processes for making enterprise-wide decisions regarding information holdings. It provides the capability to determine ownership and data standard adoption processes, to address data integrity, to define processes for business-process development, and to establish a mechanism for arbitrating differences. The benefits are that it decreases data duplication, improves cost effectiveness of data sharing throughout the enterprise, and increases data quality Data Management."

This section discusses data management topics governed through the structures and processes described in the MES Governance Plan. This section describes cross agency data stewards, Agency data stewards, and Agency data owners. In the MES Governance plan, each of these roles is a specialized type of subject matter expert. Data stewards and data owners perform important roles of identifying and communicating issues requiring decisions and implementing decisions resulting from the MES Governance processes. The MES Governance processes and structure support the MES Data Management and Data Governance. The following discussion about MES Data Governance does not imply a different or additional governance structure for the MES Program.

4.1.1 USE OF MES GOVERNANCE PROCESSES

MES Data Governance is a specific implementation and use of the MES Program governance processes and framework defined in the Medicaid Enterprise System Governance Plan. For the MES Program, Data Governance refers to the overall management of the availability, usability, integrity, and security of the data available in the enterprise.

Although the Agency creates and controls a vast amount of Medicaid data, it also relies on other state agencies to provide critical Medicaid data. Conversely, the Agency also shares a large amount of Medicaid data with other state agencies. Data governance is critical to help data consistency and quality across all state agencies and all business processes. MES Data Governance describes the use of the MES Governance Plan to make and implement MES Data Management.

The MES Governance provides a tiered structure and processes that provide leadership, guidance, decision making and overall direction for the MES Program.

The MES Technology Governance (MTG) is the primary governance entity for data management subjects within MES Governance. In addition to technical representation, the MTG has broad business representation to ensure representation of the voice of the business.

A MTG data governance workgroup will assign specific data management subjects as needed. Additional business resources can be added to the MTG if concerns of insufficient business representation occur. The MTG can escalate or coordinate with other MES governance entities as needed. The MES Technology Governance supports both planning and control functions.

Planning function responsibilities include:

- Understand agency strategic direction and data needs

- Understand data management strategy

- Approve data governance roles and responsibilities

- Engage data stewards and data owners as subject matter experts

- Approve data principles and standards

Control functions responsibilities include:

- Make decisions related to Agency data management and operations

- Review data quality scorecards

4.1.2 GOVERNANCE PROCESS

Data governance and data strategy decision making processes are like MES Technology Domain standards setting processes. For this reason, the Agency and SEAS Vendor will leverage the processes and tools used for MES Technology domain standards.

The governance of MES Technology domain standards follows a defined process to communicate, support vendors, assess compliance, and report compliance to MES Technology domain standards. A summary of the defined process is that:

- The Agency MES Technology Governance (MTG) governance entity makes and approves technology decisions related to the technology assets of the Agency. While data is a business asset supported by technology, the data governance process aligns with MES Technology Governance. For this reason, MTG is the entry point for data governance decisions.

- The SEAS Vendor researches, advises, and prepares materials for MTG governance approval. The SEAS Vendor develops communication materials, implements the communication processes, provides MES Project Vendor support, conducts compliance assessments and reports compliance to the Agency.

- The Agency MES Technical Domain Lead directs the SEAS Vendor and authorizes the release of communications, providing vendor support, conducting compliance assessments and reporting compliance to the Agency.

The document that describes the complete process is accessible via hyperlink below to the document on the MES Repository:

- SEAS NH T-6 Technology Standards Communication, Support, Compliance and Compliance Reporting Procedures

4.1.3 LIST OF SUBJECT AREAS FOR GOVERNANCE

As per the MITA 3.0 Information Capability Matrix (ICM) guidelines, there should be data governance policies and procedures in place within the Agency and across data consumers.

Table 6-1. MITA ICM

Business Area Title					
	Level 1	*Level 2*	*Level 3*	*Level 4*	*Level 5*
Data Management Strategy (DMS)					
Does business area have governance of data management?	No data governance implemented.	Implementation of internal policy and procedures to promote data governance, data stewards, data owners, and data policy.	Adoption of governance process and structure to promote trusted data governance, data stewards, data owners, data policy, and controls redundancy within intrastate.	Participation in governance, stewardship, and management process with regional agencies to promote sharing of Medicaid resources.	Participation in governance, stewardship, and management process with Centers for Medicare & Medicaid Services (CMS) and other national agencies and groups to promote sharing of Medicaid resources.

Exhibit 4-1: MITA 3.0 Information Capability Matrix (ICM) Guidelines

Exhibit 4-2: Subject Areas that Require Data Governance Process provides example subject areas that require data standards and policies.

SUBJECT AREA	EXAMPLES
Recipient	Enrollment, Eligibility, Demographics, Benefits Plan
Provider	Enrollment, Eligibility, Demographics, Specialty, Benefits Plan
Contractor	Outreach, Medicaid Contracts, Medicaid Solicitation
Plan	Goals, Plans, Policies, Benefits, Set-Rate
Care	Plan of Care, Diagnosis, Screening, Case Management, Transition Plan

SUBJECT AREA	EXAMPLES
Financials	Receivables, Expenditures, Payments, Rebates, Adjustments
Performance	Measurement Metrics, Expected Standards, Physician Quality Reporting System (PQRS)
Operations	Claim Activity, Invoice, Payment Liability, Remittance
Business Relationship	Outreach, Communications
Technology Standards	Governance details available in T-6 Technology
Data Standards	Governance details available in T-3 Data Standards
Security	Governance details available in T-8 Enterprise Data Security Plan

Exhibit 4-2: Subject Areas that Require Data Governance Process

4.1.4 DATA STANDARDS

Data standards are the rules that enable interoperability, data sharing and data consistency. The SEAS T-3 Data Standards deliverable defines MES data standards and data standards governance process.

4.1.5 EXECUTIVE SPONSOR

An executive sponsor is a "Chief"-level executive (e.g. CIO, CTO, CDO) who is an active participant and advocate for Agency Data Management and Data Stewardship processes, disciplines, and activities. The executive sponsor is responsible for staffing, funding, business prioritization, and cross functional collaboration.

4.1.6 DATA GOVERNANCE LEADER

An individual, independent of data ownership, who coordinates tasks for Data Stewards, helps communicate decisions made by Data Stewards to relevant stakeholders, drives ongoing data auditing and metrics that assess program success and ROI, and is the primary point of escalation to the Executive Sponsor and MTG. The workgroup lead of the MTG data governance workgroup will serve in this role.

4.1.7 AGENCY DATA STEWARDS

Agency Data Stewards are functional subject matter experts who provide valuable input into the MTG data management and data stewardship process. The responsibilities of Data Stewards include:

- Serving as the business function expert for the data for a specific business segment

- Defining data quality metrics and validation rules

- Defining and maintaining an enterprise data glossary

- Defining and maintaining business metadata including items such as data lineage, business definitions, validation rules, etc.

- Creating and submitting new policies or standards, as needed, to confirm appropriate data entry, usage, and transformation

- Maintaining business logic for data movement, transformations, and defining validation approaches

- Coordinating with other Data Stewards for enterprise data stewardship needs

4.1.8 CROSS AGENCY DATA STEWARDS

Cross Agency Data Stewards, while not part of MTG, are representative business experts from agencies other than AHCA who provide valuable input to the data management and data stewardship processes. Their responsibilities include understanding strategic priorities of the shared Medicaid data and relating to their agency specific processes, tools, participating in defining rules, standards, data quality metrics, and championing data quality within their agency. The holistic enterprise approach seeks to reduce duplicated data using integration technologies to provide real time information exchange and service reuse. Cross system information exchange relies on integration governance processes. When integration enables information exchange across organizational boundaries there is a need for governance and coordination. The strategy recognizes that as information crosses boundaries there are data stewardship activities to optimize business operations as new and modified information exchanges evolve.

4.1.9 AGENCY DATA OWNERS

Agency Data Owners are individuals who are responsible for supporting and implementing data governance and best practices for data elements. Data Owners will need to be identified and appointed for each area within the Agency. Data Owner responsibilities include:

- Serve as the business resource responsible for the day-to-day management of data including custody, safekeeping, usage, and integration with the business functional area

- Support Data Stewards on current and future data governance activities

- Identify, support, and prepare documentation including data integration, sources of data origin and downstream systems

- Maintaining knowledge of data structures to assist in the creation of entity relationship diagrams in a format that is acceptable to the Agency

- Help define and update data quality standards and technical rules as needed to improve data quality and consistent use of data

- Implement data confidentiality and security requirements

4.2 DATA GOVERNANCE RACI CHART

Exhibit 4-3: Data Governance RACI presents a Responsible, Accountable, Consulted, and Informed (RACI) table for an initial set of MES data governance activities.

DATA GOVERNANCE ACTIVITY	AGENCY / MTG GOVERNANCE	SEAS VENDOR	ISIP VENDOR	CROSS AGENCY DATA STEWARDS	AGENCY DATA STEWARDS	AGENCY DATA OWNERS	DATA ARCHITECT / DBA
Define data quality metrics and validation rules	A	C	I	I	R	C	I
Define and maintain business metadata, including items such as data lineage, business definitions, validation rules etc.	A	C	I	I	R	C	I
Create or update new standards	C	A	I	I	R	C	I
Coordinate with other data stewards for cross agency enterprise data stewardship needs	C	C	A	C	R	C	I
Inform the Agency on strategic data needs - business perspectives	C	A	C	I	R	C	I
Review and approve data architecture	C	C	A	I	C	C	R
Address technical matters as related to data	I	C	A	I	C	C	R

DATA GOVERNANCE ACTIVITY	AGENCY / MTG GOVERNANCE	SEAS VENDOR	ISIP VENDOR	CROSS AGENCY DATA STEWARDS	AGENCY DATA STEWARDS	AGENCY DATA OWNERS	DATA ARCHITECT / DBA
Document and update technical metadata	I	C	A	I	C	C	R
Prepare documentation including data integration, sources of origin and downstream systems from a technical perspective.	I	A	C	I	C	R	C
Define and maintain an enterprise data glossary	A	C	I	I	R	C	I
Update data model in the enterprise repository	I	A	C	I	C	C	R
Define or revise existing data sharing architecture	C	C	A	I	C	C	R
Provide access to data	A	C	R	I	C	I	I
Implement data confidentiality and security requirement	C	C	A	I	C	R	C
Inform agency's strategic data needs - technical perspectives	C	C	A	I	C	C	R
Approve agency strategic direction on data topics	AR	C	C	I	I	I	I
Develop and maintain data management strategy	R	A	C	I	C	C	I
Establish data governance roles and responsibilities and appoint data stewards and custodians	R	A	C	I	I	I	I
Maintain Data Management Governance Charter	R	A	C	I	I	I	I

Exhibit 4-3: Data Governance RACI

SECTION 5   COMMON DATA ARCHITECTURE

The MES Common Data Architecture is a pillar of the overall MES Data Management Framework that includes Data Modeling and Data Service and Integration. These capabilities support the Data Sharing Architecture pillar of the MES Data Management Framework.

Per MITA version 3.0 Part II Chapter 2, "The Common Data Architecture establishes standard data-management procedures for the data models. The data architecture provides specific guidelines regarding data documentation, data-sharing development and use applicable to both structured and unstructured data, and management of metadata of all types. These guidelines are in place so the SMA defines data entities and attributes, data models, and relationships to convey the overall meaning and use of Medicaid data and information. Common data architecture improves the consistency in the development and deployment of systems, increases the data quality, decreases the complexity of the Extract, Transform, and Load (ETL) process, decreases resource expenditures, and increases performance. The Seven Standards and Conditions, State Self-Assessment (SS-A), MITA Maturity Models (MMM), Concept of Operations (COO), and Business Process Models (BPM) provide a foundation for common data architecture."

5.1 COMMON DATA ARCHITECTURE

The Agency's future state data architecture is the foundation for realizing the Agency's Data Management Strategy. Common data architecture is about managing the data elements found in common in more than one of the key systems. Common data elements are sometimes referred to as master data or reference data elements.
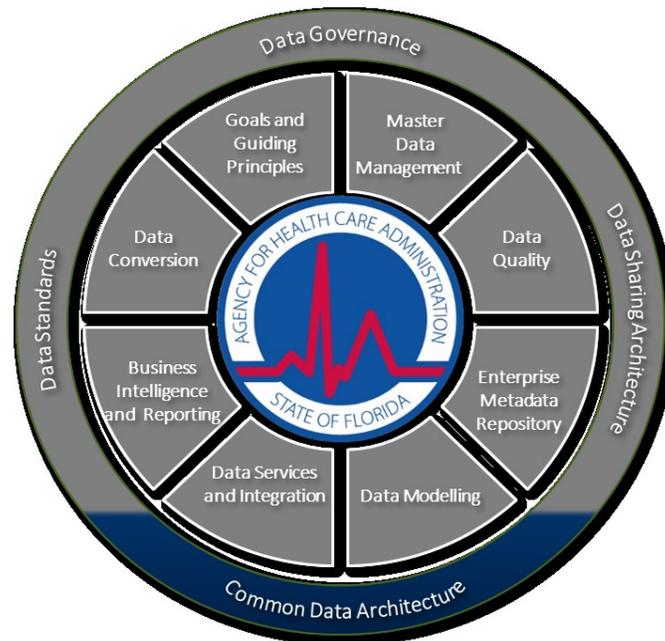
Exhibit 5-1: Data Management Capability Framework – Common Data Architecture

This section presents MES direction on strategic topics in the Common Data Architecture pillar of the Data Management Capability Framework. Topics addressed in this release of the Data Management Strategy focus on master data management, master data services, and related capabilities required for implementation of the MES Infrastructure and future MES module implementations. Section 6 addresses the data modeling aspects of common data management.

Section 8 Data Sharing Architecture describes data sharing guidelines, data architecture scenarios, pros and cons for the Medicaid enterprise.

T-4 Technical Management Strategy Section 3.6 Cloud Computing contains the strategy for Cloud computing including data management considerations for adoption.

Strategic Topic 5-1: Common Data Architecture describes the Agency overall common data architecture strategy.

COMMON DATA ARCHITECTURE	Current	2018	TIMELINE 2020	2022	2025
Continue without a common data repository and shared data services. Common data is replicated across multiple repositories.	X	X			
Establish a common data architecture strategy implemented as a master data repository.		X			
Each Module to use common master data repository and data services.			X	->	
Analysis	Common data elements are those data elements that are generally found across multiple repositories throughout the enterprise. Examples of common data are Name, Address, Phone Number, and E-mail. Other examples of master data entities include Provider, Recipient, and Health Plans. In the current environment, the Agency lacks a strategy to manage common data elements or master data elements. In the future state, Agency will have a common data repository for some of the data elements that are considered critical and yet commonly used, there will also be common data services to create and access those data elements. It is expected that common data elements will be identified by the business and implemented as such by EDW and module vendors. Common data elements will be implemented with the following considerations: ▪ Identify the common data elements ▪ Develop the common data model ▪ Implement into a schema that can be accessed by other modules or module implementing vendors ▪ Provide common data services to access these common data elements				

Strategic Topic 5-1: Common Data Architecture
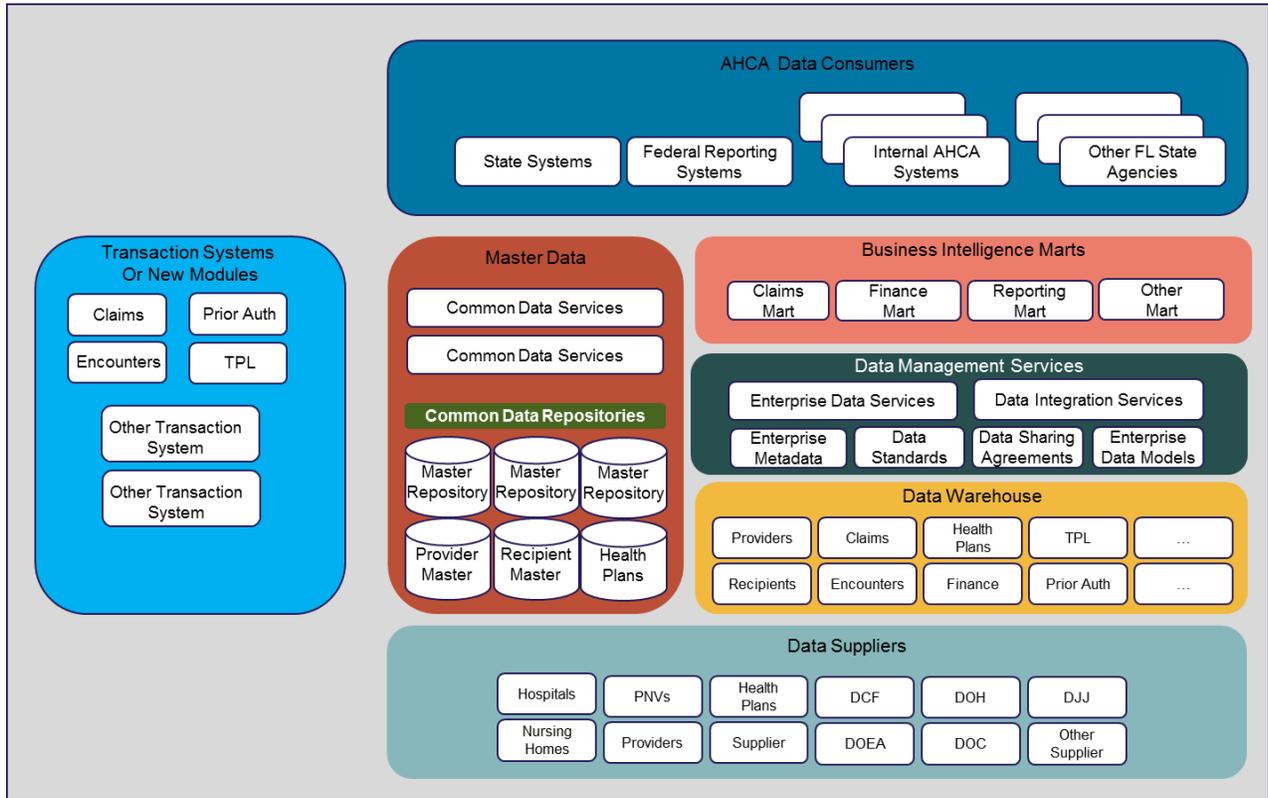
Exhibit 5-2: Future State Data Architecture is the future state data environment that has Common Data Architecture as a foundational building block.

AHCA Data Consumers

State Systems | Federal Reporting Systems | Internal AHCA Systems | Other FL State Agencies

Transaction Systems Or New Modules

Claims | Prior Auth
Encounters | TPL
Other Transaction System
Other Transaction System

Master Data

Common Data Services
Common Data Services

Common Data Repositories

Master Repository | Master Repository | Master Repository
Provider Master | Recipient Master | Health Plans

Business Intelligence Marts

Claims Mart | Finance Mart | Reporting Mart | Other Mart

Data Management Services

Enterprise Data Services | Data Integration Services
Enterprise Metadata | Data Standards | Data Sharing Agreements | Enterprise Data Models

Data Warehouse

Providers | Claims | Health Plans | TPL | …
Recipients | Encounters | Finance | Prior Auth | …

Data Suppliers

Hospitals | PNVs | Health Plans | DCF | DOH | DJJ
Nursing Homes | Providers | Supplier | DOEA | DOC | Other Supplier

Exhibit 5-2: Future State Data Architecture

5.1.1 COMPONENT: MASTER DATA MANAGEMENT

Master Data Management (MDM) is the disciplined process in which IT engages business experts in the development and maintenance of consistent and accurate definitions of an enterprise's most critical information. An MDM methodology focuses on eliminating redundancy, inconsistency, and fragmentation by having a single, synchronized, comprehensive, authoritative source of master information. For example, almost every entity within the Agency and outside the Agency uses Provider and Recipient data. These data types affect every operational and analytical system in its daily work.

Example of master data entities include:

- **Provider Master** – A person or organization enrolled in the Florida Medicaid program that provides healthcare services.

- **Recipient (Master Person Index)** – A person who is eligible for assistance in healthcare related services in accordance with the Florida Medicaid program.

- **Health Plan** – An organization that provides coverage to plan recipients for designated services at a fixed premium.
- **Claim** – A request for Medicaid to pay for healthcare services.
- **Encounter** – Information about individual healthcare related services provided by a health plan.

5.1.2 COMPONENT: MASTER DATA SERVICES

Many business components require complete, timely, authoritative, and quality master data information. Master data services are a set of core data services that enable interaction with the underlying master data objects. Examples of master data services include creating provider record or search for provider details by unique identifier(s).

The Agency's DMS will implement Master Data Management to enable reliable and consistent data with an integrated 360˚ information view of (at least) providers, recipients, and health plans.

Strategic Topic 5-2: 360˚ view of Recipient describes the Agency strategy for providing a 360˚ view of recipient data.

360˚ VIEW OF RECIPIENT DATA	Current	2018	TIMELINE 2020	2022	2025
Ad hoc - No strategy	X	X			
MPI Modules are operational including data sharing strategy.			ODS, RDS, NoSQL evaluation	->	
360˚ view of information as per cross agency data governance.				Align to multi-agency data federation strategy	->
Analysis	Currently, the Agency lacks a unified strategy for supporting the 360˚ view of a recipient. The future state strategy is to provide a 360˚ view of recipient through a combination of MPI and advanced data matching techniques. This same pattern can be extended to other areas beyond recipient in the future.				

Strategic Topic 5-2: 360˚ view of Recipient

5.1.3 MASTER DATA MANAGEMENT CAPABILITY STRATEGIES

This section briefly explains strategic direction for MES use of various master data management technologies and techniques.

- **Master Data Management** – The MES data management strategy is to establish registry style master person and master organization indexes that link records that have common identities. This strategy accepts that the enterprise can reduce, but never eliminate the creation of master data records and transaction records for people and organizations that share the same identity. The MES strategy is to centralize identity matching logic to automate identity matching using probabilistic and deterministic identity matching methods. The matching engine will perform real-time linking of identity records. The identity records will maintain unique keys to source systems for records linked to an identity. Pulling the collection of data using the identity keys from all linked identity records will provide access to a 360˚ view of information about a person or organization. When the master person and master organization indexes are established, they can be used to reduce the creation of new duplicate master records and simplify reuse of information already collected about a person or organization.

- **Master Person Index** – The MES strategy for the Master Person Index is to procure and configure COTS software to perform real-time high volume identity matching and linking. The approach creates and maintains links for any type of record about a person. The Master Person Index will ultimately include millions of entries as records in the many systems in the Agency, State of Florida and other systems are linked.

- **Master Organization Index** – The MES strategy for the Master Organization Index parallels the strategy for the Master Person Index. The Master Organization Index will use COTS MDM software to identify identity linkages. Organization linkages can be found with more complex identity matching rules that consider ownership, name changes, and tax identity linkages.

- **Establish Unique Identifiers** – The MES strategy related to defining unique identifiers will be to minimize investment in trying to assign the same record identifier to all instances of records for a person. Instead of struggling with insurmountable challenges including trying to change history, the MES strategy will invest in the matching technology and master person and master organization indexes and make applications recognize and process considering records for all linked identities.

- **Match and Merge for Identity Resolution** – The Master person index and Master organization index will be the source of truth of matching records for the same person or organization. The MES strategy recognizes that merging records of matching identities is extremely complex and expensive. The MES strategy may support simple merges of information but intends to focus investments on presenting and operating on aggregated sets of information assembled from all matching identity records. The ODS will be the source of record for person and organization data used in source of truth master person and master organization matching.

- **Publish-Subscribe** – The MES strategy is to use publish-subscribe technology to push communication and notice of updates or business transaction events that occur in one system to other systems for all of the linked identity records in the Master Person Index or Master Organization Index. Systems that receive notices from publish-subscribe services will make system specific decisions on how to act on the information. Some systems may auto apply changes to their systems, others may initiate processes (e.g. address confirmation or eligibility screening) on event notification and some may manually process or address at next system interaction with the person.

- **Golden Record** – The MES strategy will be to invest in the registry style master person index and master organization index and provide a holistic view of the aggregate information of all linked identity records. In health and human service systems, the pursuit of golden record strategies is usually difficult.

- **Manual Matching** – While the techniques mentioned above will resolve many matching and merging use cases, they will not resolve 100% of the cases. The Agency will need to continue to have operational personnel and processes to address manual matching scenarios. As part of the manual matching process and as manual matching cases are resolved, they should be coded accordingly so going forward they are automatically matched rather than manually matched.

Strategic Topic 5-3: Golden Record Strategy describes the Agency strategy related to the concept of implementing a Golden Record approach.

GOLDEN RECORD	Current	2018	TIMELINE 2020	2022	2025
Ad hoc - FMMIS is used as the Golden Source for operational data. Exception cases exist where additional external data is used.	X	X			
Source of truth services using advanced data matching techniques to align data into single view			X	->	
Cross agency data governance with data federation				Explore / Pilot / Limited	->

GOLDEN RECORD	Current	2018	TIMELINE 2020	2022	2025
Analysis	Currently, FMMIS collects a large quantity of data from other agencies, providers and organizations external to the Agency. FMMIS is thus acting as a single source of truth for some Agency data. In the future state, AHCA in collaboration with other agencies should explore a cross agency data governance model with data federation to determine source of truth. Each agency should have the accountability to keep their master data clean, maintain its data model on an easily accessible location, maintain a data dictionary, and maintain metadata as per the cross-agency data governance process. While this is the target for 2022, FMMIS should continue to be the golden source for operational data until module vendors implement advanced data matching techniques to establish the golden source of truth if the operational data exists in multiple sources.				

Strategic Topic 5-3: Golden Record Strategy

5.2 MASTER DATA OBJECTS

As per MITA 3.0, there are nine different subject areas, as outlined below, where master data management should be applied. These subject areas all exist within Agency systems.

SUBJECT AREA	EXAMPLES
Recipient	Enrollment, Eligibility, Demographics, Benefits Plan
Provider	Enrollment, Eligibility, Demographics, Specialty, Benefits Plan
Contractor	Outreach, Medicaid Contracts, Medicaid Solicitation
Plan	Goals, Plans, Policies, Benefits, Set-Rate
Care	Plan of Care, Diagnosis, Screening, Case Management, Transition Plan
Financials	Receivables, Expenditures, Payments, Rebates, Adjustments
Performance	Measurement Metrics, Expected Standards, Physician Quality Reporting System (PQRS)
Operations	Claim Activity, Invoice, Payment Liability, Remittance
Business Relationship	Outreach, Communications

Exhibit 5-3: Master Data Objects

Focus areas for master data management are Provider, Recipient (Member), and Plan. Once master data management processes and procedures have been successfully established, they should then be extended to other areas within the Agency as listed above.

5.3 COMMON DATA ARCHITECTURE GUIDELINES

Common Data Architecture will adhere the guidelines as listed in Section 3.6 - Data Management Guiding Principles.

5.3.1 MASTER DATA MANAGEMENT TOOLS

Master Data Management tool evaluation and selection will occur as part of the Integration Services and Integration Platform procurement. The SEAS Vendor defined Master Data Management requirements to include in the competitive solicitation. Some of the industry leading Master Data Management tools that support the Agency's common data architecture are:

- Informatica Siperian
- IBM MDM
- Oracle MDM
- Talend MDM
- SAS MDM

SECTION 6   ENTERPRISE MODELING

An enterprise data model is a holistic view of the data consumed and produced across an entire organization and represents a single integrated definition of data, irrespective of systems or applications. Enterprise data modeling is the process of creating, maintaining, and sharing the data models necessary to support the information systems within an organization. Most organizations create three types of data models:

- Conceptual Data Models
- Logical Data Models
- Physical Data Models

A description of each type of data modeling follows in Section 6.2.4 below.

The following SEAS Technology Deliverables include additional MES Data Modeling content:

- T-2 Information Architecture Documentation – defines conceptual and logical data models
- T-3 Data Standards – defines the data dictionary and data standards

6.1   DATA MODELING STRATEGY OVERVIEW

This section provides an overview of the strategic direction for MES Data Modeling.

6.1.1   TYPES OF DATA MODELING

The main categories of data to be modeled for the MES program are listed below. These categories would typically be present in both logical and physical data models.

- **Operational Data Modeling** – This is modeling of the operational data including customer records and transaction records used by applications and systems (e.g. claims, recipient, provider data)

- **Information Exchange Data Modeling** – This modeling is for information passed between systems or services. The information in an information exchange is transitory and may not permanently reside as operational data.

- **Analytic Data Modeling** – This modeling is for information in the data warehouse, data marts, reporting data stores used for dashboards, reporting, ad hoc inquiry, and predictive analysis.

- **Rule and Policy Data Modeling** – This modeling is for information about business rules and policy data used by rules engines, data validation services and validation engines.

- **Experience Data Modeling** – This modeling is of the interaction lifecycle of recipients and providers. This type of modeling supports collection and analysis of behavior determinant data.

6.1.2 ROLES IN PERFORMING DATA MODELING

The SEAS Vendor will establish the data modeling strategy, processes and tools to support modeling of enterprise capabilities and solution components. MES Project Vendors use the established processes and tools to develop, provide and maintain the content specific to each MES Project or module.

6.1.3 DATA MODELING PRIORITIES

The MES Program needs definition, structure, tools, policies and processes for all categories of data modeling. The sequence to establish data modeling processes and capabilities is:

- **Integration Information Exchange** - Establish the data modeling process and content to enable information exchange using the Integration Platform:
  › Integration platform standard message structure and framework
  › Service registry and repository data
  › System User Identity and Access Data
  › Data used to populate the Master Person Index and Master Organization Index
  › Publish subscribe data framework
  › Rules Engine data
- **Operational Information Exchange** - Establish the data modeling process and content for exchange of operational data. This modeling is the overall information exchange canonical model and information exchange development methodology. NIEM based modeling would be the default data modeling strategy.
- **Operational Data Store** - Establish data modeling processes and capabilities to support establishment of the Operational Data Store. This modeling defines the subject area-based information model that will support Data-as-a-Service (DaaS) access methods.
- **Analytic Data** - Establish Data modeling processes and capabilities to model Analytic data for the Reporting Data Store, Analytic Data Store and Data Marts to support the Enterprise Data Warehouse Project. This work will be done in parallel with the establishing modeling processes and capabilities for the Operational Data Store.
- **Rule and Policy Data** – Establish data modeling or electronic business rules, data edits, data validations, data translations, and policy.
- **Experience Data** – Establish data modeling processes and capabilities to capture and store recipient and provider experience data.

6.2 ENTERPRISE DATA MODELING CAPABILITIES

This section describes the key Agency data modeling capabilities to successfully create, maintain, and share data models across the Agency. While each of these capabilities is important, the Agency's maturity in each will vary and grow over time. That difference among capabilities and growth path is expected and not unique to the Agency.

6.2.1 BUSINESS DRIVEN

Data models are a representation of the data requirements needed to support the business processes and systems for an organization. Data models are business-driven. When considering the needs of the business it is imperative to consider the full scope of the organization's requirements. In the context of the Agency, that includes the operational needs of the Agency to process Medicaid claims and the many downstream data requirements within Medicaid and outside Medicaid.

6.2.2 ACCESS TO MODELS THROUGH A SHARED REPOSITORY

Providing authorized shared access to those individuals directly interacting with data models is critical to confirm technical alignment and understanding across the Agency. This is a key capability that will support data modeling of data exchanges, interfaces and modular integrations. All data models (conceptual, logical, and physical) will be controlled through a central shared repository with appropriate access and authorization based on roles within the Agency. Typical roles who will need access to the data model repository are IT Managers, Technical Architects, Data Stewards, Programmers, Testers, Business Analysts, and Technical Writers. Most top-rated enterprise modeling tools provide some type of shared repository capability.

6.2.3 UPDATED MODELS

Data models are only beneficial when they are accurate and represent the current state of the organization and its data assets. In fact, often outdated and inaccurate data models can do more harm than good. Section 6.3 Enterprise Data Modeling Governance covers the process, roles, and responsibilities to confirm data models are accurately maintained in a timely manner.

6.2.4 TIERED DATA MODELS

The Agency will maintain three distinctly different levels of data models: Conceptual, Logical and Physical data models. These three different data models each serve different and unique needs within the Agency and represent the logical progression from business requirements to the physical database. The three data models are:

- **Conceptual Data Model** – A high-level depiction of the Agency's information needs and assets. Conceptual data models describe key data and relationships in a

technology and application agnostic way and are often used with stakeholders during initial requirements gathering.

- **Logical Data Model** – A diagram that represents the entities, attributes, and relationships involved in a specific business function. A conceptual data model will typically decompose into multiple logical data models. Logical data models are the basis for the creation of the physical data model. Depending on purpose, logical data models can be either application agnostic or specific to an application.

- **Physical Data Model** – A database technology-specific implementation of the logical data model. Physical data models contain tables, columns, keys, relationships, constraints, and any other necessary data objects required to implement a specific system or application.

6.2.5 COMMON ENTERPRISE DATA MODELING TOOL

While it is technically possible for an organization to use multiple enterprise data modeling tools, the MES strategy is to standardize on a COTS enterprise data modeling tool across the Agency. This will aid in consistency across data assets and systems and reduce unnecessary waste created when divisions within the Agency or external vendors use their individual data modeling tool of choice. Use of a common tool promotes efficiency and accuracy when modeling data exchanges, interfaces, and modular integrations. Section 6.4 Enterprise Data Modeling Tools describes the key capabilities required in an enterprise data modeling tool. The MES strategic direction is to standardize on the use of Erwin as its enterprise data modeling tool.

6.2.6 LINK TO ENTERPRISE DATA DICTIONARY

T-3 Data Standards, Section 3 Data Definition contains information about building and maintaining an Agency data dictionary. Linking enterprise data models to the data dictionary provides exponential benefit in terms of the understanding and usability of the data models. When complete, the data dictionary in conjunction with the business glossary becomes the business vocabulary that provides invaluable context to the data models. Section 6.4 Enterprise Data Modeling Tools describes the key capabilities required in an enterprise data modeling tool one of which is "business vocabulary definition mapped to usage within models".

6.2.7 MODELING STANDARDS FRAMEWORK

The creation, consistent use, and enforcement of standards for data modeling is critical to confirm users who interact with data models can do so in a meaningful way. This promotes efficiency and accuracy when modeling data exchanges, interfaces and modular integrations. These standards include both industry best practices for data modeling and Agency specific standards. All of the modeling standards listed below will be followed. Modeling standards will include:

- Model and model object naming

- Creation of an enterprise glossary

- Primary key, foreign key, and secondary indices guidelines

- Data sensitivity classification

- Data type guidelines

- Stored procedure guidelines

- View guidelines

- Versioning using a [major version].[minor version] technique

  › Increment major version when changes are made that break backward compatibility (e.g. adding or removing a required element, changing the type of an attribute, changing the meaning of an existing component)

  › Increment minor version when changes are made that do not break backward compatibility (e.g. adding optional elements or optional content)

6.3   ENTERPRISE DATA MODELING GOVERNANCE

Data modeling governance, not to be confused with data governance, is an important element of Agency data modeling to confirm the availability, usability, integrity, and security of the data model assets. The following recommendations are categorized according to the four principles mentioned above (availability, usability, integrity, security). Section 6.4 Enterprise Data Modeling Tools describes the key capabilities required in an enterprise data modeling tool, many of which align to these four data modeling governance categories. This reinforces the importance of tool selection as an enterprise data modeling tool not only is the mechanism to create, update and share data models but also is the tool that helps enforce the Agency's data modeling governance processes.

6.3.1   AVAILABILITY

The MES Data Modeling strategy provides appropriate access to enterprise data models through a central repository, role-based authorization, and data model file formats.

The Agency will maintain a central repository, enabled though the enterprise data modeling tool, of all enterprise data models. This central repository will be organized in a way that allows easy navigation and searching. Access to the central data model repository will be role based and allow for standard CRUD (create, read, update, delete) operations based on each user's role. The central repository will also be organized with a section or folder for archived data models, which will serve as a historical reference. Contained within this section of the repository are all the data models that are no longer in use or active. Within the data model central repository, as enabled by the enterprise data modeling tool, models will be available in a variety of file formats to meet each user's need (e.g. tool native format, image, PDF, UML). Models will be versioned according to Section 6.2.7 Modeling Standards Framework.

6.3.2 USABILITY

To confirm the usability of data models within the Agency, they must be complete, accurate and updated on a timely basis. Prior to any new project starting, the enterprise data models will be evaluated for usability and to determine what, if any, areas of data ownership, management, and stewardship will be impacted. Updates to data models will be performed in conjunction with all system changes that affect either the conceptual, logical, or physical data model. No database change, regardless of size or scope, should be done without first updating the corresponding data models. Data model change documentation is required to be included as part of the Change Advisory Board (CAB) process, or Agency equivalent pre-release certification process. In cases of a Severity 1 incident where a critical system is down and services restoration is top priority, database changes made to restore system functionality must be reflected accordingly in the data model and approved as part of a post-Severity 1 retrospective process. These processes maintain the accuracy and completeness of all data models in the central repository.

6.3.3 INTEGRITY

Data model integrity begins with having a set of modeling standards as described in Section 6.2 Enterprise Data Modeling Capabilities under the Standards Framework. Governance to enforce these standards will be in place and will be a layered approach including the use of standard data modeling templates, managerial review and approval of data model changes, requiring data model change documentation as part of CAB approval and periodic audits of the central data model repository conducted by a Data Steward.

Key data modeling governance roles are:

- **Data Architect / DBA** – Responsible for creating, updating, or deleting data models per Agency standards and guidelines. Can be at some or all model levels (conceptual, logical, and physical)

- **Enterprise Data Architect** – Responsible for reviewing and approving data model changes and helping establish data modeling standards, guidelines, and templates.

- **Change Manager** – Responsible for enforcing data modeling standards and guidelines through the CAB process.

- **Data Steward** – Responsible for helping establish data modeling governance and enforcing compliance to standards and guidelines through regular audits.

6.3.4 SECURITY

Data models in all forms (conceptual, logical, and physical) contain highly sensitive information and should be controlled accordingly. A data model in the hands of the wrong person can easily result in a security breach through techniques like SQL Injection. Because of the sensitivity of data models, access to the central data model repository will be controlled with

role-based security and allow for standard CRUD (create, read, update, delete) operations based on each user's role. Typical users who need access are IT Managers, Technical Architects, Data Stewards, Programmers, Testers, Business Analysts, and Technical Writers. From a governance perspective, a timely process will be in place to remove access from the central data modeling repository when a user leaves the Agency or transitions to another role that no longer requires access to data models. In addition, a quarterly audit will be conducted to identify individual outliers where access should be revoked.

6.4 ENTERPRISE DATA MODELING TOOLS

Based on AHCA's diverse data landscape including significant operational data (e.g., providers, recipients, claims, and encounters), robust data warehousing needs throughout the Agency, and future modular implementation approach, the Agency will standardize on a common enterprise class data modeling tool. Tool standardization helps promote commonality across the data landscape and provides unity between implementation vendors as modules are developed and deployed.

Below is a list of key modeling capabilities the MES Program is considering when evaluating and adopting the enterprise data modeling tool. The capabilities in this list are Core Modeling capabilities and Management and Collaboration capabilities. Both capability categories are important when considering enterprise data modeling tools.

6.4.1 CORE MODELING CAPABILITIES

- Support for three levels of data modeling (conceptual, logical, and physical)
- Ability to generate new models from existing models
- Model validation and error checking against industry standards and best practices
- Support for technology-specific object types in physical data models
- Management of dependencies between models and between model objects
- Reuse of models and model objects, including model patterns
- Business vocabulary definition mapped to usage within models
- Ability to define tables, columns, defaults, constraints, indexes, stored procedures, triggers, and relationships
- Ability to import and export data models
- Ability to merge data models
- Ability to compare database models at the physical level to map the differences
- Ability to reverse engineer and document models from existing databases

- Ability to create a database by generating SQL Data Definition Language (DDL) code or by creating a live database using the appropriate SQL dialect for the target database platform
- Ability to generate test data from a physical data model
- Support for multiple database platforms (e.g. SQL Server, Oracle)

6.4.2 MANAGEMENT AND COLLABORATION CAPABILITIES

- Export information from model objects for publication in other formats such as HTML, Word, Excel, PDF, and XML.
- Support for a shared location, either file system or database, for the storage of and controlled access to data models
- Capability to resolve potentially conflicting changes made by different DBAs or Data Architects
- Share reference models and common business rules via a common repository
- Ability to modify best practices to fit individual corporate standards
- Ability to define a business vocabulary or glossary

Currently, the Agency is using Erwin as the data modeling tool for operational data in FMMIS. The AHCA IT team uses Visual Studio Entity Designer as their data modeling tool. Many quality enterprise data modeling tools exist in the marketplace which meets Agency needs. Some of the top-rated data modeling tools are:

- Erwin Data Modeler
- IBM InfoSphere Data Architect
- IDERA ER/Studio
- Sybase PowerDesigner

6.4.3 MODELING TOOL RECOMMENDATION

The MES modeling tool strategy is to standardize and use Erwin as its enterprise data modeling tool for the following reasons:

- Erwin is the CMS recommended data modeling tool per the CMS.gov website under Tool Use.
- Erwin is already used as the data modeling tool for operational data in FMMIS
- The existing established vendor relationship can be leveraged in future license negotiation and acquisition

- Erwin meets the data modeling functional needs of the Agency. While other data modeling tools above also meet the Agency's functional data modeling needs, Erwin is the Agency's incumbent tool and has historically and continues to be one of the top-rated data modeling tools in the marketplace

- Erwin is platform and database technology agnostic which will provide needed flexibility as modules are being developed and deployed

- Erwin should be the data modeling tool used exclusively within the Agency. Vendors implementing modules may use data modeling tools other than Erwin as long as those tools offer data models in native Erwin, or Erwin compatible, formats to allow importing into the central data model repository

- Erwin Data Modeler is one product in a broader suite of products that may also prove useful to the Agency (e.g. Erwin Data Governance, Erwin Enterprise Architect and Erwin DM NoSQL)

Strategic Topic 6-1: Enterprise Data Modeling Tool Strategy describes the recommended enterprise data modeling tool strategy.

DATA MODELING TOOLS	Current	2018	TIMELINE 2020	2022	2025
Erwin tool and native Erwin format	FMMIS	FMMIS	Agency ALL / Module Vendors	->	
Visual Studio Entity Designer	AHCA IT	AHCA IT			
Erwin Compatible Model Format			Module Vendors / Other Agencies	->	
Analysis	Erwin is the CMS recommended data modeling tool and meets the data modeling functional needs of the Agency. While other data modeling tools also meet the Agency's functional data modeling needs, Erwin is the Agency's incumbent tool and has historically and continues to be one of the top-rated data modeling tools in the marketplace. Erwin is platform and database technology agnostic which will provide needed flexibility, as modules are being developed and deployed. Erwin Data Modeler is one product in a broader suite of products that may also prove useful to the Agency (e.g. Erwin Data Governance, Erwin Enterprise Architect and Erwin DM NoSQL). Existing established Erwin vendor relationship can be leveraged in future license negotiation and acquisition.				

Strategic Topic 6-1: Enterprise Data Modeling Tool Strategy

6.5   ENTERPRISE DATA MODELING SCOPE

Given the Agency's approach to modular implementation and Integration Services, consideration should be given to the scope of data that is appropriate to model. The following three categories exist:

- Data that should be modeled:
    - › Operational data including ODS and Reporting Data Store (RDS)
    - › Data warehouse, Data marts, and Reporting data
    - › Any Agency owned and controlled data that is stored in a database
    - › Data and Metadata used in the exchange of information through the Integration Services and Integration Platform (ISIP)
- Data that should not be modeled:
    - › Modular implementations where vendor COTS solutions bundle system functionality with proprietary data schemas and solutions
    - › Software as a Service solutions
    - › Platform as a Service solutions
- Data that may be modeled:
    - › Data from external sources or agencies used by AHCA but not owned or controlled by AHCA

SECTION 7   ENTERPRISE METADATA REPOSITORY

Metadata is descriptive data or information about other data. Metadata management is the process of administering and using metadata to inform key business decisions. Metadata management defines policies and procedures to access, integrate, share, analyze and maintain information across the organization.

The Agency does not currently have a formal metadata management tool or mature metadata management processes. iTrace contains some metadata, AHCA IT has some metadata in extended properties in MS SQL Server, and there are spreadsheets across the Agency that contains metadata for key data elements. Most of the Agency does not use a formal metadata repository, tool, or process.

7.1   ENTERPRISE METADATA REPOSITORY CAPABILITIES

This section describes the key metadata repository capabilities that will exist within the Agency to successfully document, maintain, and use metadata throughout the Agency.

The metadata repository serves as the central element and enabler for metadata management. Most organizations have a significant volume of metadata, but few are disciplined in collecting and publishing metadata in a central repository. The metadata repository is used to document, manage, and share metadata. Data Owners and Data Stewards are most often the providers and managers of content in the metadata repository. Users of metadata within the repository are typically data consumers, business analysts, system architects and designers, system testers, and system support staff.

7.1.1   DATA LIFECYCLE

The data lifecycle describes in narrative and graphical form how data moves through an organization and systems. The lifecycle includes the origin of data, how and where data moves over time, how data is transformed as it moves through the organization and systems and how and when data is archived.

7.1.2   GLOSSARY

A comprehensive list of business terms and technical terms used within the Agency. The glossary will include the list of terms, their definition, and any linkages or relationships between the terms.

7.1.3   METADATA DISCOVERY

The Agency, like many organizations, has metadata spread throughout the enterprise in many forms and formats. Metadata discovery is the process and associated tools to extract metadata

from existing sources and incorporate it in the metadata repository. These sources include data models, databases, BI tools, Word and Excel documents, policies, and procedures.

7.1.4 BUSINESS RULES MANAGEMENT

Business rules must be documented and include their linkage to both data and metadata. This can be accomplished within the scope of the metadata repository or a separate business rules engine. If documented in a separate business rules engine, it is imperative that the appropriate linkage exists within the metadata repository to connect the business rules to the corresponding metadata.

7.2 ENTERPRISE METADATA MANAGEMENT PROCESSES

Enterprise metadata management is not as much a technical or data problem as it is a business, people, and process problem. This section will address the key processes and roles and responsibilities that need to be in place to effectively and efficiently maintain and leverage an enterprise metadata repository.

The following processes must be in place and governed for the Agency to create, maintain, and use an enterprise metadata repository.

7.2.1 CRITICAL DATA ASSET IDENTIFICATION PROCESS

Identifying critical data assets identifies important data, and corresponding metadata reduces the risk of overlooked or missing data in the metadata repository. The process will use a top down approach to identify critical data asset first and subsequently review successively lower levels of data. A goal of critical data asset identification is to eliminate small, insignificant differences between divisions, departments or groups and align on a common metadata definition through consensus of data stakeholders. As the Agency matures in metadata management processes and repository comprehensiveness, the scope of the metadata repository should be extended to include those partners, organizations, and agencies which AHCA does business with on a regular basis. This extension should only be considered once the metadata repository process and content are robust from an internal Agency standpoint.

7.2.2 CHANGE AND APPROVAL PROCESS

The Metadata Manager will prescribe a process, enforced by Data Owners, for how and when additions and changes will be made in the metadata repository. This process must be widely communicated within the Agency. The purpose of this process is not to be restrictive but to guard the integrity of the content in the metadata repository. Changes to metadata should be made in conjunction with corresponding system changes or be made according to the Agency's standard release cycle. Data Owners are the gatekeepers for all metadata changes and therefore must approve all changes. A critical component of the change and approval process is content versioning. All changes to content within the metadata repository must be version

controlled so users and systems can reference a specific metadata definition by version and understand the changes across versions.

7.2.3 STANDARDS ENFORCEMENT

Metadata standards enforcement is the responsibility of Data Stewards. Standards enforcement occurs through a multi-tiered approach. Primarily, standards enforcement will occur by Data Stewards on a regular metadata repository update cycle. It is the responsibility of Data Stewards to review and provide feedback on the compliance to metadata standards before making changes to the metadata repository. Additionally, Data Stewards must conduct a regular (every six months recommended) audit of the metadata repository to evaluate standards compliance. To the extent the Agency metadata management tool of choice supports the configuration and enforcements of repository standards they should be fully leveraged.

7.2.4 ACCESS AND AUTHORIZATION

Metadata is an extremely sensitive data asset and must be controlled accordingly. The Agency must appropriately control access and authorization to the metadata repository through role-based security. A timely process must be in place to remove access from the metadata repository when a user leaves the Agency or transitions to another role that no longer requires access to the repository.

Key roles and responsibilities related to metadata management are:

- **Metadata Manager** – Overall responsibility for the operation of the metadata repository and metadata management governance process.

- **Data Steward** – Responsible for enforcing compliance to metadata standards and managing metadata governance metrics.

- **Data Owner** – Responsible for the creation and maintenance of metadata within the metadata repository. This includes metadata they create themselves and metadata additions or changes submitted by others that they approve.

- **Data Consumer** – Anyone, either internal or external to the Agency, who uses, accesses, or consumes in any way data related to the Agency. These consumers, whether human or system, should be given appropriate accesses to the metadata for the data they consume.

7.3 ENTERPRISE METADATA MANAGEMENT TOOLS

There are many quality metadata management tools in the marketplace that will meet AHCA's metadata management needs. When evaluating and selecting a metadata management tool it is critical to consider the capabilities listed in Section 7.1 Enterprise Metadata Repository Capabilities. The Agency does not currently have a single mature metadata management tool

in place. Some of the top-rated and most widely used metadata management tools in the marketplace are:

- Collibra Data Governance Center, Collibra Catalog and Collibra Connect
- IBM InfoSphere Information Governance Catalog
- Informatica Metadata Manager, Business Glossary, Axon and Enterprise Information Catalog
- Oracle Enterprise Metadata Management
- Smartlogic Semaphore

The future state strategy is to store, maintain and disseminate metadata from a single COTS based metadata management tool. While any of the metadata management tools listed above would represent a significant step forward for the Agency in the area of metadata management, the Informatica suite of tools provide the most robust metadata management solution and have the broadest adoption in the marketplace.

Strategic Topic 7-1: Enterprise Metadata Management Tool Strategy describes the recommended enterprise metadata management tool strategy.

METADATA MANAGEMENT TOOLS	Current	2018	TIMELINE 2020	2022	2025
iTrace	X	X	X (Read Only)		
Spreadsheets and Documents	X	X			
Database Extended Properties	X	X			
SharePoint	X	X			
Informatica Exports (e.g. PDF, Excel)			Non-Informatica Users	->	
Informatica			X	->	

METADATA MANAGEMENT TOOLS	Current	2018	TIMELINE 2020	2022	2025
Analysis	Currently, the Agency maintains metadata in a combination of tools and systems (e.g. iTrace, Spreadsheets and Documents, Database Extended Properties). The future state strategy is to store, maintain and disseminate metadata from a single COTS based metadata management tool. Data management is critical to the Agency with ever increasing demands and value placed on data quality, accessibility, and transparency, hence the need for an Agency wide COTS metadata management tool. Among the top metadata management tool vendors, Informatica Metadata Manager is considered an industry leader and recommended for Agency wide metadata management.				

Strategic Topic 7-1: Enterprise Metadata Management Tool Strategy

SECTION 8   DATA SHARING ARCHITECTURE

The Data Sharing Architecture section:

- Provides a high-level overview of the MES Data Sharing architecture

- Describes the major components of the data sharing architecture and strategic topic direction related to specific data sharing components. These major components are the building blocks that promote modularity, reuse, easier integration, and less duplication.

- Describes acceptable data sharing approaches and data sharing reference architecture

- Lists recommended tools for data sharing

8.1 HIGH-LEVEL TO-BE CONCEPTUAL DATA SHARING ARCHITECTURE DIAGRAM

AHCA's Data Sharing Architecture promotes both MITA and Agency data management goals of interoperability and data availability. In the current environment, most Agency data is available for data sharing. While FMMIS and DSS are the backbone operational and decision support systems, there is significant healthcare data in other data repositories and formats (e.g. SQL Server for business intelligence and reporting, licensure information in VERSA, hospital discharge data, financial data in a variety of financial data stores). The Agency uses all datasets to support internal processes or to share data with other non-Agency systems.
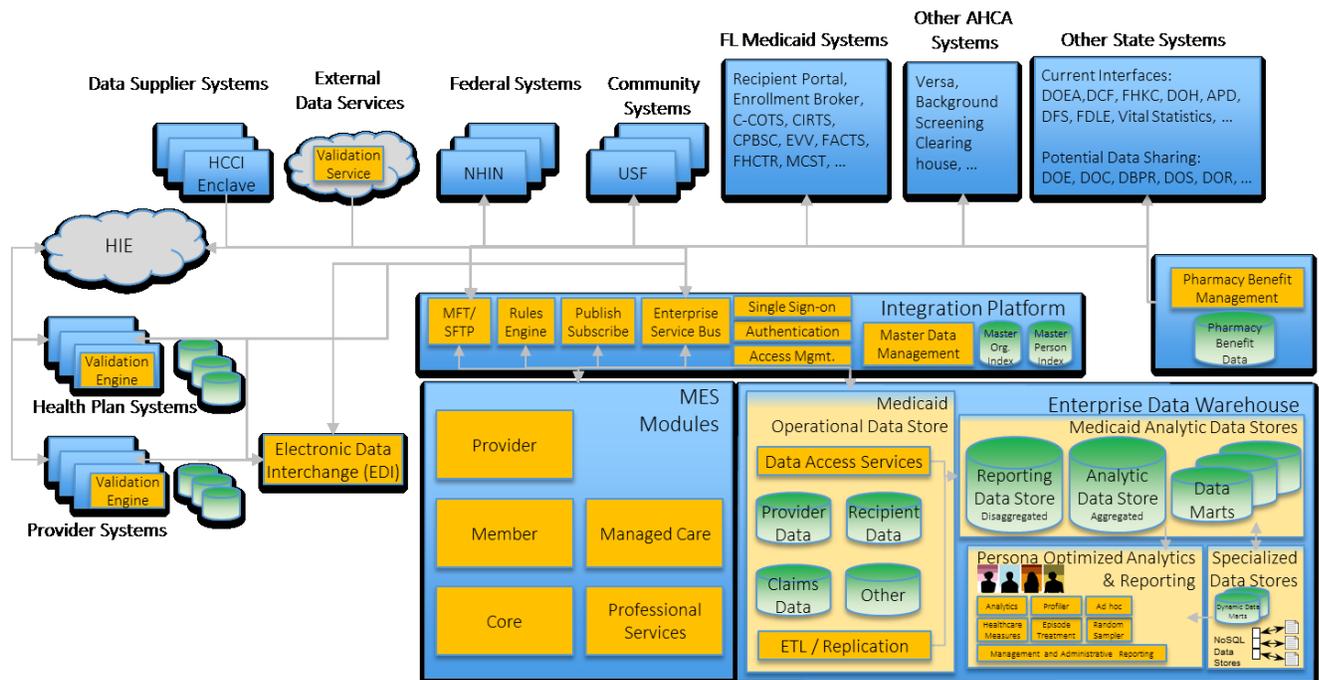
Exhibit 8-1: High-Level To-Be Conceptual Data Sharing Architecture Diagram

8.2 DATA SHARING ARCHITECTURE COMPONENTS

The components of the data sharing architecture establish new integration and interoperability capabilities and improve the existing data sharing capabilities. Each of the components contributes to improved data sharing. Each component adds incremental features. In total, the fully integrated solution of all components produces value in excess of the sum of each individual component.

8.2.1 INTEGRATION PLATFORM COMPONENTS

Section 3.1.2 Data Management Strategy Vision Enabling Capabilities provides a summary description of each component of the overall High-level Data Architecture Conceptual Diagram. Additionally, the Integration Services and Integration Platform (ISIP) procurement document specifies detailed requirements for each component that enables and improves data sharing.

8.2.2 DATA SHARING ARCHITECTURE DATA STORES

Four primary data stores will contain and manage MES data.

- Operational Data Store
- Reporting Data Stores
- Data Warehouse / Analytic Data Store
- Data Marts

The MES Program strategy is to design each Data Store to aid data sharing and use of the data stores as a single source of the truth. The combination of underlying technology and data design will allow these data stores to be used for their defined business purpose.

8.2.2.1 OPERATIONAL DATA STORE (ODS)

The Operational Data Store is the implementation of a data strategy that establishes a single source of record for transactional data. Other sources of record exist both inside and outside the Agency and will be identified and documented as part of ISIP, EDW and modular implementation design. As described in section 5.1.3 Master Data Management Data Management Capability Strategies Match and Merge for Identity Resolution, the ODS will be the source of record for person and organization data used in source of truth master person and master organization matching. Data in the ODS is independent of a specific application or system. For this reason after migration to the ODS, the Agency will be able to replace a module from one vendor with modules from another vendor. Applications access data in the ODS using data service or API calls as opposed to passing SQL language directly to a proprietary

database. AHCA IT is in the adoption process of using this pattern for Agency internal IT projects.

The MES strategy to migrate to use of the ODS is to incrementally migrate information while maintaining a real time data replication from the ODS to FMMIS. This approach minimizes disruption to FMMIS during the transition to the ODS.

Strategic Topic 8-1: ODS Build and Transition Strategy describes the Agency's Operational Data Store (ODS) build and transition strategy.

ODS BUILD AND TRANSITION	Current	2018	TIMELINE 2020	2022	2025
FMMIS Database is the ODS	X	X			
Incremental migration of data to the ODS (e.g. by module). Data in ODS is not replicated with FMMIS			Initial module (e.g. provider)		
Big bang migration of all data to the ODS with data replication with FMMIS. FMMIS continues to use its database, not the ODS.				Remaining modules	
Big bang migration to the ODS without data replication.					
Analysis	The strategy of an incremental migration is lower risk than a single big-bang ODS implementation. The direction to replicate data between FMMIS and the ODS reduces the need for modifications to FMMIS systems and processes during modular implementations. Balance of ODS data will be enabled in ODS as part of wave 2 ('Remaining Modules') ahead of the actual module implementations.				

Strategic Topic 8-1: ODS Build and Transition Strategy

8.2.2.2   REPORTING DATA STORE (RDS)

The MES data management strategy is to have dashboards and reports access a Reporting Data Store (RDS) for MES data. A reporting store is an optimized data store for real time

reporting, dashboards and ad hoc queries. Use of a reporting data store reduces contention and performance impact to the ODS.

Strategic Topic 8-2: RDS Refresh Frequency Strategy describes how often data from the ODS is updated in the Agency's Reporting Data Store (RDS).

RDS REFRESH FREQUENCY	Current	2018	TIMELINE 2020	2022	2025
Weekly					
Daily					
Periodic scheduled (e.g. hourly, multiple per day)					
Near real-time (continuous update using queue, log shipping, replication, etc.)	Non-FMMIS	Non-FMMIS	ODS / RDS	->	
Analysis	Currently, the Agency lacks a Reporting Data Store (RDS) for Medicaid processing outside of the operational reports available in FMMIS. The future state strategy will provide near real-time operational updates to the RDS to support operational data reporting, improved data quality, and down-stream system synchronization. This will result in reduced load on the ODS by separating operational reporting load from transaction processing load. The RDS will also be optimized for operational reporting needs while the ODS will be optimized for transaction processing needs.				

Strategic Topic 8-2: RDS Refresh Frequency Strategy

8.2.2.3 DATA WAREHOUSE / ANALYTIC DATA STORE

Data warehousing is a set of techniques and software to enable the collection of data from operational systems, the integration and harmonization of that data into a centralized database and then the analysis, visualization and tracking of key performance indicators. Data warehouse, data marts, operational data store (ODS), and reporting data store (RDS) are the backbone components for achieving business intelligence and enabling the ability to gain insight into the business. The business value delivered includes answers to decision support questions about providers, recipients, claims, financials, and more.

8.2.2.3.1 DATA WAREHOUSE CAPABILITIES

This section describes the key data warehouse capabilities that must exist within the Agency to successfully create, maintain and use the MES Enterprise Data Warehouse (EDW). While each of these capabilities is important, the Agency's maturity in each will vary and grow over time. That difference among capabilities and growth path is expected and not unique to the Agency.

- Data Integration
    - › Integrate with multiple Agency owned data sources and data stores external to the Agency
    - › Integrate with existing and future business intelligence and data analytics tools to enhance the Agency's decision-making activities around fraud, waste, and abuse detection and prevention
    - › Perform Extract, Transform and Load activities (ETL)
    - › Provide near real-time data replication
- Data Management:
    - › Store all data and metadata required for the administration and operation of the Medicaid program
    - › Scale as required to meet the growing data needs of the MES, preserving historical information as necessary
    - › Manage disparate data sets across the MES
- Data Access
    - › Provide an architecture for decision makers to access data for enterprise-wide data analysis and reporting
    - › Provide data housing and data integration capabilities for data to be shared across system boundaries
    - › Enable multiple types of data services (e.g. elemental and composite)
    - › Support transaction processing across operational systems involving relatively small volumes of data routed as transactions through the ESB
- Metadata
    - › Prepare and provide metadata including a data dictionary and business glossary
    - › Identify critical data elements where metadata is required
    - › Support data lineage report views
- Data Quality
    - › Support data enhancement, data validation, and data confidence
    - › Identify and resolve data conflicts and data quality defects

- Data Standards
  - › Support Agency specific data exchange standards for integration and sharing

8.2.2.3.2 DATA WAREHOUSE ARCHITECTURE STRATEGY

The MES Data Strategy is to implement a centralized data warehouse for analytic and reporting data needs that are not supported by the RDS. This strategy recognizes the challenges with data sharing across agencies and centralizes control. As the Agency's data warehouse implementation matures, data sharing with external agencies can be explored through federation.

Strategic Topic 8-3: Data Warehouse Architecture Strategy describes the Agency's data warehouse architecture strategy.

DATA WAREHOUSE ARCHITECTURE	Current	2018	TIMELINE 2020	2022	2025
Ad hoc - Multiple distributed repositories as warehouses	DSS, SQL Server, Others	DSS, SQL Server, Others	DSS, SQL Server, Others		
Central warehouse for all Agency Medicaid analytic and reporting data needs			EDW	->	
Central warehouse for all Agency analytic and reporting data needs				EDW	->
Central warehouse for multi-agency (e.g. AHCA, DCF, DOH) analytic and reporting data needs					
Federated warehouse across Agency boundaries (e.g. Medicaid, HQA)					

DATA WAREHOUSE ARCHITECTURE	Current	2018	TIMELINE 2020	2022	2025
Federated warehouse for multi-agency analytic and reporting needs enabled through semantic layer					
Hybrid of centralized warehouse within the Agency and federated with external agencies, providers, and health plans				Explore / Pilot / Limited	->
Analysis	The future state strategy is to implement a centralized data warehouse for all analytic and reporting data needs within the Agency. This strategy recognizes the challenges with data sharing across agencies and centralizes control. As the Agency's EDW implementation matures, data sharing with external agencies can be explored through federation. Centralized: Traditional replication model commonly adopted for business intelligence. In this model, data is consolidated and blended into single physical repository. A typical centralized data warehouse repository would contain metadata and business rule information to aid maintenance. It will also include data quality and data governance specific processes. The biggest benefits of a centralized model include: 1) high degree of control of the data owned by the Agency, 2) ability to enforce governance, and 3) high performance and easier data maintenance. Drawbacks include: 1) more expensive to provision and support 2) longer lead times before model can be 'seeded' and used, and 3) more latency in distribution. Federated: In this model, the repository reduces/eliminates the need for data replication (e.g., data aggregation functions can be virtualized). Federated repositories maintain metadata and transformation logic to create standardized and customized views meeting various data consumers' needs and aid 'rapid marts' creation from authoritative sources to satisfy changing business requirements. Benefits include: 1) no data replication 2) lower cost to build and support 3) allows for data to be made available in real-time or near real-time, and 4) improved ability to respond to new business requirements. Drawbacks include: 1) poor performance for complex virtual views 2) more complex to maintain metadata currency 3) requires stronger governance and stakeholder commitments, and 4) aggregation, summarization, and derivation is difficult.				

Strategic Topic 8-3: Data Warehouse Architecture Strategy

8.2.2.3.3 DATA WAREHOUSE TECHNOLOGY TYPE

Strategic Topic 8-4: Data Warehouse Technology Type Strategy describes the Agency's data warehouse technology type strategy.

DATA WAREHOUSE TECHNOLOGY TYPE	Current	2018	TIMELINE 2020	2022	2025
Relational for warehouse	DSS / SQL Server	DSS / SQL Server	EDW	->	
NoSQL (e.g. Hadoop) for Advanced Analyst persona		Pilot (early 2019)	->		
NoSQL (e.g. Hadoop) for entire warehouse with relational data marts				Evaluate	

DATA WAREHOUSE TECHNOLOGY TYPE	Current	2018	TIMELINE 2020	2022	2025
Analysis	Currently, the Agency's core data warehouse (DSS) resides in an Oracle relational database. Other warehouses also exist (e.g. Analytic SQL Server), which are also relational. Unstructured data is primarily stored on file systems and accessed and combined with relational data on an ad-hoc basis. The future state strategy is to implement a relational data warehouse and data marts to meet the majority of the reporting and analytic persona needs of the Agency. In addition to the relational data warehouse and data marts, a NoSQL database will be implemented to meet the storage, indexing, searching and reuse needs of the Agency's unstructured data. This capability is primarily targeted to support the needs of the Advanced Analyst persona. Longer-term, an assessment of using NoSQL for the entire warehouse, with relational data marts, will be conducted. Relational: The benefits of a relational data warehouse include: extensive resources are available and familiar with RDBMS (Relational Database Management System); current Agency systems, processes and tools are designed and built for RDBMS; RDBMS technology is optimized to meet heavy query and reporting needs; BI tool landscape is mature. The drawbacks include: performance ceilings exist and can be costly to address, scaling requires expensive hardware (e.g. horizontal clustering) for both performance and redundancy, minimal support within RDBMS exists for unstructured data in a sustainable manner. NoSQL: The benefits of a NoSQL data warehouse include ability to scale on low cost commodity hardware, built-in native redundancy, distributed computing model for performance and redundancy, support for unstructured data via a schemaless design, strong solution for real-time analytics when combining varying data types and sources, increasing interoperability tools between RDBMS and NoSQL. The drawbacks are less familiar technology vs. RDBMS, limited but growing BI tool/connector landscape, Agency systems, processes and tools are not targeted for NoSQL. NoSQL resources for implementation and support are also considerably higher, anywhere from 150% to 170%, than similar relational SQL resources.				

Strategic Topic 8-4: Data Warehouse Technology Type Strategy

8.2.2.4 DATA MARTS

Data Marts are datasets that are organized and optimized for use by a specific business unit and persona type combination. MES will implement two types of data marts: fixed and dynamic. Fixed data marts (e.g. MAR, federal reporting, division specific analysis) retain their organization and structure for ongoing analysis and are refreshed on regular intervals, determined by the data needs in the data mart, from the EDW. Dynamic data marts watch user

behavior and usage and dynamically adapt the data mart dimensions, facts, relationships, indexes, etc. to best meet user needs. Section 3.10 Business Intelligence (BI) describes Business Intelligence personas, which are the target data mart user community.

Strategic Topic 8-5: Data Mart Reporting Strategy describes the Agency's Data Mart reporting strategy.

RDS REFRESH FREQUENCY	Current	2018	TIMELINE 2020	2022	2025
Operations reporting in FMMIS. Ad hoc Data Marts created as a reactionary measure. Multiple copies of reporting/analytic data	X	X			
Reporting Data Store (RDS) for operational reporting			X	->	
Reporting and Analytics Data in Warehouse			X	->	
Relational Marts & OLAP by key business areas configured by persona			X	->	
Analysis	In the future state, the Agency is expected to have a single Operational data store (ODS) and a corresponding reporting data store (RDS). The recommendation for the Agency is to have relational data marts and Online Analytical Processing (OLAP) models by key business areas configured by persona. The warehouse will address the detailed reporting and analytics data requirements whereas the RDS will address the operational reporting requirements. The ODS should be strictly used by the modules and should support the operational transaction processing requirements.				

Strategic Topic 8-5: Data Mart Reporting Strategy

8.2.2.5 UNSTRUCTURED DATA

In the future state, an Enterprise Data Warehouse (EDW) and Enterprise Service Bus (ESB) are core enablers to AHCA's Data Sharing Architecture for structured data. Unstructured data will be stored, indexed and used via a NoSQL database.

Currently, there is not an Agency wide strategy for managing non-traditional, often unstructured, data (e.g. behavioral, customer experience, photo, sensor, genetic, etc.). The methods for obtaining, storing and using these data are ad-hoc across the Agency.

The future state strategy is to implement a NoSQL database to support the storage, management and use of unstructured data. The use of unstructured data via the NoSQL database will largely focus on the Advanced Analyst persona and supplement the relational data warehouse and data marts.

Strategic Topic 8-6: Data Warehouse Architecture Strategy describes the Agency's support for Agency analytics that bring in non-traditional data types.

NON-TRADITIONAL DATA TYPES	Current	2018	TIMELINE 2020	2022	2025
Ad hoc - No Agency wide strategy. Non-traditional data landscape is not well defined or supported. Multiple provider systems and processes for storing and consuming data exist.	X	X			
Collect and store behavioral, photo, genetic, etc. data as structured data					
Collect and store behavioral, photo, genetic, etc. data in NoSQL database and start meaningful analytics (e.g. 360˚ view of recipient, fraud analysis)			Advanced Analyst persona	->	

NON-TRADITIONAL DATA TYPES	Current	2018	TIMELINE 2020	2022	2025
Analysis	Currently, there is not an Agency wide strategy for managing non-traditional, often unstructured, data (e.g. behavioral, customer experience, photo, sensor, genetic, etc.). The methods for obtaining, storing and using these data are ad-hoc across the Agency. The future state strategy is to implement a NoSQL database to support the storage, management and use of unstructured data (e.g. narrative reports, investigative and court reports, call session recordings, email content, documents, behavioral, customer experience, photo, sensor, genetic, etc.). The use of unstructured data via the NoSQL database will largely focus on the Advanced Analyst persona and supplement the relational data warehouse and data marts.				

Strategic Topic 8-6: Data Warehouse Architecture Strategy

8.3 DATA SERVICES AND INTEGRATION

The Data Services and Integration is the component of the Data Management Framework that provides the framework for data access and data sharing. This section describes the MES strategic direction on the use of data services, integration using data sharing reference architectures, and the use of data sharing agreements.
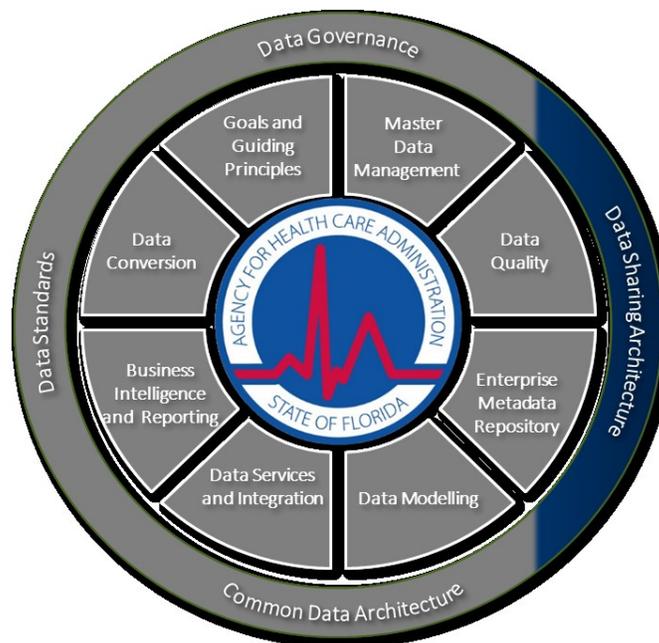

Exhibit 8-2: Data Management Framework – Data Sharing Architecture

8.3.1 DATA SERVICES

As part of the DMS, reusability is a key guiding principle and hence the Agency and its vendors will design and develop data management solutions with reusable data services. These data services will be registered and published as a library of reusable services. Currently, the Agency's use of data services is limited to select systems within AHCA IT where the Microsoft Entity Framework is being used.

The future state strategy is to leverage data services at varying levels of granularity within and across Agency systems. Data service granularity will span from elemental data services to module specific data services to composite cross module data services.

Strategic Topic 8-7: Data Service Granularity Strategy describes the Agency's data services granularity strategy.

DATA SERVICE GRANULARITY	Current	2018	TIMELINE 2020	2022	2025
No Data Services - Direct SQL	X	X			
Elemental data services (e.g. CRUD operations). No public visibility.	Non-FMMIS	Non-FMMIS	X	->	
Module specific data services. Local module visibility.	Non-FMMIS	Non-FMMIS	X	->	
Composite data services for reuse across modules. Public visibility.			X	->	
Analysis	Currently, the Agency's use of data services is limited to select systems within AHCA IT where the Microsoft Entity Framework and Oracle platform are being used. The future state strategy is to leverage data services at varying levels of granularity within and across Agency systems. Data service granularity will span from elemental data services to module specific data services to composite cross module data services.				

Strategic Topic 8-7: Data Service Granularity Strategy

8.3.2 INTEGRATION - DATA SHARING REFERENCE ARCHITECTURES

Data Integration is the process of "exchanging" data between two processes, systems, or organizations. The Enterprise Service Bus enables data integration methods on a real-time basis. Extract Transform Load (ETL) supports data integration through batch services.

Strategic Topic 8-8: System Data Integration Strategy describes the Agency's acceptable methods for sharing data between and across system boundaries.

SYSTEM DATA INTEGRATION	Current	2018	TIMELINE 2020	2022	2025
Custom master file replication	X	X			
Custom batch via ETL	SFTP	SFTP	Large Volume Only	->	
Custom integration to common cross system repository					
Direct DB to DB enabled via COTS			Preferred	->	
Publish-Subscribe			Preferred	->	
Real-time event based via services / ESB	Point to Point	Point to Point	Preferred for Small Data Volumes	->	
Analysis	Currently, the majority of the Agency's cross system data integration is enabled through the exchange of files, custom data transformation processes, SFTP, and some point-to-point system integration. The future state strategy will provide multiple methods for system data integration based on the size, type and immediacy needs of the data being exchanged. The Agency's data landscape is diverse in terms of data types and volume. The MES data integration strategy implemented as part of ISIP will provide multiple, right sized, solutions for data integration. This targeted approach improves data timeliness and relevancy.				

Strategic Topic 8-8: System Data Integration Strategy

8.3.2.1 DATA INTEGRATION THROUGH ENTERPRISE SERVICE BUS

The Enterprise Service Bus (ESB) provides a communication system where software applications interact in a service-oriented architecture (SOA) that allows sharing of data between producers and consumers. The SOA is an organization-wide, shared, reusable service model used by all applications integrated into the ESB. Software applications integrated in this manner are called services. The ESB performs message management, service authorization and access control, availability management, usage and cost accounting, and

service coordination for complex orchestration of services. The ESB decouples the network design from the underlying platform and allows the Agency greater data integration and sharing possibilities with modern technologies across multiple vendors. Data sharing through the ESB will largely be used for smaller, real-time, data exchanges. As the Agency moves from a monolithic MMIS system to a modular MMIS system the ESB becomes the communication broker and enabler between modules and systems, whether internal to the Agency or external to the Agency.

8.3.2.2 DATA INTEGRATION THROUGH ETL

Another integration approach is to leverage ETL. ETL covers a process of how data is loaded from the source system into a target system. ETL will most often be used when the Agency needs to move large volumes of data between systems. This could be nightly batch processes or the ingestion of large data sets from external Agency partners.

There are several techniques to perform an Extract process. The most popular are:

- Through Publish-Subscribe notification where source system(s) publish a notification that a data record has been modified at the source and describe the nature of the change. The Subscribing system monitors events of a certain type and then responds accordingly to those events.

- Incremental extract where the source system(s) perform a change data capture on a period basis to capture the records that have been modified and provide an extract of such records.

- Full extracts where source system(s) provide a full copy of all available data (e.g. complete copy of data for disaster recovery purposes).

8.3.2.3 USE OF SOCIAL DETERMINANTS OF HEALTH DATA

Strategic Topic 8-9: Social Determinates of Health Strategy describes the Agency's strategy for use of social determinates of health.

SOCIAL DETERMINATES OF HEALTH	Current	2018	TIMELINE 2020	2022	2025
Ad hoc - No strategy. Individuals desiring to use social determinates of health for analysis are on their own to find solutions.	X				

SOCIAL DETERMINATES OF HEALTH	Current	2018	TIMELINE 2020	2022	2025
Begin supporting common business processes from Ad hoc sources		X			
Develop sourcing and data sharing strategy and plan			X		
Collect and store social determinates of health data for advanced analytics			NoSQL or NoSQL / Relational Hybrid	->	
Analysis	Currently, there is not an Agency wide strategy for the sourcing, sharing and use of social determinates of health data. Select individuals / departments are using social determinates of health data but processes are entirely Ad hoc. The future state strategy is to develop and execute a plan for sourcing, storing, and using social determinates of health data across the Agency. Given the structured, semi-structured and unstructured nature of this broad category of data, either a NoSQL database or hybrid NoSQL / Relational database will be required.				

Strategic Topic 8-9: Social Determinates of Health Strategy

8.3.3 BLOCKCHAIN

Blockchain is a distributed technology allowing users to record data through an ever-expanding list of "blocks". Each data record is written into blocks, time-stamped and connected to the block that came before it, setting up a blockchain. Furthermore, each block cannot be modified, deleted or otherwise changed.

Rather than a central database, the Blockchain data records are distributed and shared across networks with credentialed users able to add to but not delete or alter the transaction log.

In healthcare, a recipient's complete medical history is spread across many different systems and organizations. This creates both security challenges as well as challenges getting a 360° view of a patient's health. By design, Blockchain addresses both of these issues through its distributed ledger and ability to link data together. For example, Blockchain could play a role in the future from enrollment to claims processing, however, it is too early for the Agency to adopt Blockchain technology for the following reasons:

- Blockchain technology is in its infancy with few implementations in healthcare

- With Blockchain being an open technology platform, securing health care information and ensuring HIPAA compliance and interoperability will be challenging

- Most vendors and vendor tools used in the industry are yet to fully adopt Blockchain technology for their products

- Since Blockchain technology is new, implementation and operation/maintenance costs could be significantly higher

Strategic Topic 8-10: Blockchain Strategy describes the Agency's Blockchain adoption strategy.

BLOCK CHAIN	TIMELINE				
	Current	2018	2020	2022	2025
Do not adopt Blockchain					
Wait for further penetration in healthcare					
Identify initial Agency use case and pilot				Evaluate / Pilot	->
Aggressively pursue adoption as a front running state					
Analysis	Blockchain appears to be a promising future technology for the healthcare industry. Given the newness of the technology, costs are high and healthcare implementations are few. The Agency should wait for the Blockchain technology space to mature a couple more years and then perform an initial evaluation / pilot.				

Strategic Topic 8-10: Blockchain Strategy

8.3.4 DATA SHARING AGREEMENTS

Data sharing agreements are an important method to confirm what data is shared, with whom, and how the data can be used. Such an agreement serves two purposes. First, it protects the Agency providing the data so that the data will not be misused. Second, it prevents miscommunication by the Agency as the data provider and the consumer receiving the data, by making certain that any questions about data use are addressed. Before any data is shared, both the provider and consumer should discuss data sharing and data use issues and come to a collaborative understanding that will then be documented in a data sharing agreement.

Securing data sharing agreements is a time consuming and expensive barrier to sharing data for the benefit of all Floridians. Today data sharing agreements are focused to specific types of data exchanges even specific fields. The MES Program will pursue implementing data sharing access controls in the integration layer of the MES architecture. With this enhanced security, MES may be able to execute data sharing agreements more efficiently.

MES data sharing agreements, at a minimum the following items should be included:

- **Types of Data and a Data Model** - describing how the data is organized.
- **Metadata** - describing data definition and other characteristics of the data.
- **Format and Process for Sharing Data** - describing how the data will be used by the consumer and the file format by which data will be shared with the consumer.
- **List of Data Services** - describing the technical services that can be reused for sharing the underlying data.
- **Data Standards** - describing the quality rules used to maintain the overall data quality at the source.
- **Data Security Standards** - used to maintain the confidential nature of the data.

8.4 RECOMMENDED TOOLS FOR DATA SHARING

The MES Program will acquire COTS tools for the Data Sharing Architecture components through competitive procurement. **Exhibit 8-3: Recommended Tools for Data Sharing** lists tools already licensed by the Agency or public domain and open source tools used for data sharing.

TOOL	USAGE / ROLE IN DATA SHARING
SoapUI	Testing tool for SOAP and REST APIs
Oracle Data Integrator	Oracle ETL / Data Integration & Transformation tool
Talend ETL	Open source ETL / Data Integration tool

Exhibit 8-3: Recommended Tools for Data Sharing

SECTION 9   MES SPECIFIC ADDITIONS

This section describes MES-specific additions of new functionality to the established MITA components for the enterprise. At this point in the MES Program, there are few additions to the MITA framework.

9.1   UNSTRUCTURED DATA

The Agency receives, generates, and manages unstructured data today from a variety of sources. In the current environment, the Agency uses LaserFiche document management, simple files, stored on shared drives, local workstations or in some cases in paper form. The Agency direction to manage this type of data is to use a NoSQL database that allows the Agency to manage the unstructured data through unified classification schemes. After data management processes classify unstructured data in taxonomies or ontology relationships, users can search, view, report on, modify, and analyze the unstructured data.

9.2   POLICY AUTOMATION DATA

The policy data that defines claims and encounter data edits, validations and data transformations is a special type of data that is a hybrid of configuration data, reference data and execution logic. The MES Program will emphasis enabling management of this type of data. The MES direction is to consolidate and harmonize all sources of this type of data use into the MES rules engine, public and third party validation services, and internally in plan and provider validation engines.