AHCA Florida Health Care Connections (FX)

T-2: Information Architecture Documentation

Version: 300

Date: June 28, 2022

Author: The SEAS Vendor

Submitted To: AHCA FX Program Administration Team

Revision History

DATE	VERSION	DESCRIPTION	AUTHOR
5/15/2018	001	T-2: Information Architecture initial draft version	JP Dahdouh Rich Cefola Paul Moore
6/12/2018	002	T-2: Information Architecture first revision	Rich Cefola
6/19/2018	003	T-2: Information Architecture first revision	Paul Moore
6/20/2018	100	T-2: Information Architecture approved baseline version	Sean Gibbs
7/29/2019	101	T-2: Information Architecture annual deliverable refresh	Barry McConnell
8/26/2019	200	T-2: Information Architecture approved annual deliverable refresh	Carol Williams
1/24/2020	201	T-2: Information Architecture Quarterly deliverable refresh	Joanna Redding Dave Schiller Steve Ruszkowski
2/6/2020	202	T-2: Information Architecture Quarterly deliverable refresh updated to address Agency comments	Joanna Redding Dave Schiller Steve Ruszkowski Barry McConnell
2/7/2020	225	T-2: Information Architecture approved quarterly deliverable refresh	Eric Steinkuehler
5/24/2022	226	T-2: Information Architecture Documentation Annual refresh draft version updated as follows: ▪ Updated Section 1 boilerplate language to reflect current approved standard language ▪ Updated all breadcrumb paths to reflect the FX Hub per DET #502	Carol Williams
6/28/2022	300	T-2: Information Architecture approved annual deliverable refresh	Amelia McHenry

Modifications to the approved baseline version (100) of this artifact must be made in accordance with the FX Artifact Management Standards.

Quality Review History

DATE	REVIEWER	COMMENTS
5/15/2018	Sean Gibbs	QA Submission Review
6/15/2018	Sean Gibbs	QA Submission Review
6/19/2018	Sean Gibbs	QA Submission Review
7/26/2019	Carol Williams	QA Submission Review
1/24/2020	Eric Steinkuehler	QA Submission Review
2/5/2020	Eric Steinkuehler	QA resubmission review
5/20/2022	`Carol Williams	Conducted quality review

Table of Contents

Table of Exhibits

Table of Strategic Topics

SECTION 1 INTRODUCTION

1.1 Background

The Florida Agency for Health Care Administration (AHCA or Agency) is adapting to the changing landscape of healthcare 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 is complex; it includes services, business processes, data management and processes, technical processes within the Agency, and interconnections and touchpoints with systems necessary for administration of the Florida Medicaid program that reside outside the Agency. The future of the Florida Medicaid Enterprise integration is to allow the Agency to secure services that can interoperate and communicate without relying on a common platform or technology.

The Florida Medicaid Management Information System (FMMIS) has historically been the central system within the Florida Medicaid Enterprise; functioning as the single, integrated system for claims processing and information retrieval. As the Medicaid program has grown more complex, the systems needed to support the Florida Medicaid Enterprise have grown in number and complexity.

The Medicaid Enterprise System (MES) Procurement Project was re-named Florida Health Care Connections (FX) in the summer of 2018. FX is a multi-year transformation to modernize the current Medicaid technology using a modular approach, while simultaneously improving overall Agency functionality and building better connections to other data sources and programs.

1.2 Purpose

T-2: Information Architecture Documentation provides a living document of the evolving information strategy, architecture, and data documentation to support the implementation of the modularized solution and establish the foundation for a data-centric organization. Its primary purpose is to provide guiding principles and a roadmap of an information architecture framework for the enterprise and allow assessment of the business areas' levels of maturity within the information architecture.

These components are necessary for the Agency to define the common data needs that will enable both near- and long-term future business processes of Florida's Medicaid Enterprise. Critical to the success of this Information Architecture are:

- Align information requirements with Medicaid Enterprise vision and direction
- Improve process effectiveness by aligning understanding of data and relationships
- Facilitate growth and innovation by creating a common information language aligned to standards

- Lower overall life cycle costs through standardization promoting the reduction of rework and test cases
- Enable interoperability and data sharing by aligning to industry standards and approaches

1.3 Scope Statement

The scope of this deliverable includes the information architecture, supporting components, and foundation of the Conceptual Data Model (CDM) and Logical Data Model (LDM). The information architecture provides a framework for understanding the scope and interrelationships of components involved in a data-centric organization and provides a basis for creating a roadmap of current and future projects. This deliverable also addresses the 10 business areas of the MITA business guidance and two (2) additional areas of integration and security. However, information architecture in general, and data modeling in particular, do not include business process modeling. As such, this document focuses on the business data concepts that define the Medicaid Enterprise but not the processing logic or operations applied to the data.

1.4 Goals and Objectives

The goal of this deliverable is to communicate the information architecture strategy, concepts, and components. An effective information architecture comes from understanding business concepts, objectives and constraints, context, and the requirements of the people that will use the information. The development of information architecture documentation occurs in various phases of FX projects. Modeling involves developing a broad understanding of the business before defining the detailed structure of data, and finally mapping of data to business process utilization.

A critical goal of this deliverable is the modeling of the future information architecture. This deliverable does not attempt to model the entire as-is environment. Modeling of the future information architecture will be more valuable to communicate direction to FX Project Owners.

An additional goal of this deliverable is to provide an analysis of the current information architecture maturity through the execution of the Information Capability Matrix as defined in the MITA 3.0 guidance.

Goal – Establish the MITA compliant Florida Information Architecture strategy. Achieving the following objectives will accomplish this goal:

- **Objective** – Define and outline data models for the Agency that incorporate the MITA standard for the business areas as described in the Scope Statement section.
- **Objective** – Use this deliverable as reference for future procurements as part of the Agency's modular implementation approach

Goal – Align the information architecture to enable interoperability and data sharing. Achieving the following objectives will accomplish this goal:

- **Objective –** Through discovery sessions and current state analysis, identify the critical pain points within the Agency related to data management
- **Objective –** Recommend frameworks, processes, technologies, and tools that provide a future vision for resolving recurring data management challenges
- **Objective –** Use the CDM as a tool between Agency users and IT Architecture to understand the core business concepts underlying FX
- **Objective –** Use the LDM in defining data entities, relationships, and meta-data to support the current operations and enable future changes with minimal disruption

Goal – The information architecture modeling meets the business needs.

- **Objective –** Use the CDM as a tool to communicate how information relates to the business functions

1.5 Referenced Documents

The following documents provided reference and guidance to the creation of this deliverable:

- MITA 3.0 Part II Chapter 3 Conceptual Data Model
- MITA 3.0 Part II Chapter 4 Logical Data Model
- MITA 3.0 Part II Chapter 6 Information Capability Matrix
- National Information Exchange Model (NIEM) V4 Reference Model
- HL7 Reference Model
- T-1: Data Management Strategy
- T-3: Data Standards
- An Introduction to the Government Information Factory by Bill Inmon

1.6 Strategic Topic Inventory

This document provides guidance on many data management strategy topics. In the development of FX Technology deliverables, 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 a hierarchical taxonomy based on logical groupings in areas of interest to strategic, programmatic, technology, and project management.

The Strategic Topic Inventory has features to present and communicate a variety of strategic direction options considered across the spectrum of time for a topic. 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 recommended strategy and direction for many of the data management strategy decisions that are important for FX Project stakeholders to understand.

Over the course of FX, the SEAS Vendor will continue to define and elaborate 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.

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

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

Area:	Service Delivery Offerings and Assets		**Description:**
Category:	Data Modeling		Who performs conceptual data modeling for the FX Conceptual Data Model?
Sub-Category	Conceptual, Logical, Physical Data Modeling		
Topic:	Who performs conceptual modeling		
Importance:		**Strategy Status:**	
Displaying Row:	462		

Strategic Direction		**Current**	**2018**	**2020**	**2022**	**2025**
SEAS vendor			X	->		
EDW Vendor				Coordination with SEAS Vendor	->	
Module Vendor						
TPA Vendor		FMMIS,DSS				
AHCA Systems (e.g. IT, HQA, ...)		X				

Analysis:	The SEAS vendor is accountable and contractually responsible for conceptual and logical data modeling. The SEAS vendor will coordinate with the EDW vendor to coordinate data services implementation issues and logical to physical modeling activities.

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	▪ Identify the data structure to meet business requirements within FX. ▪ Evaluate data elements within Business Areas. ▪ Maintain the Information Architecture document. ▪ Identifies the data management related technologies and processes necessary to improve FX. ▪ 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. ▪ Maintain the Agency Data Management Strategy. ▪ Support vendor procurements by providing information, extracts, and details related to the Data Management Strategy. ▪ Initiate and approve changes to names, formats, and definitions of the data models and elements including performing coordination and validation with the Data Stewards, Data Governance Leads, and Data Owners.
AHCA FX Technical Lead	▪ Provide direction and recommendation for escalated modeling definition issues where the Data Architect, Data Stewards, Data Governance Leads, and Data Owners disagree. ▪ Promulgate these definitions and structures across future system design projects, as appropriate. ▪ Coordinate 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. ▪ Communicate data management strategy to AHCA FX Leads. ▪ Support FX leadership communications to Agency executive leadership. ▪ Approve communications between the SEAS Vendor and FX Stakeholder Organizations related to FX Data Management Strategy. ▪ Coordinate data management governance processes through the FX Data Governance Workgroup.
FX Project Vendors (IS/IP, EDW, Module)	▪ Follow the strategic direction in the Data Management Strategy in proposing, discussing, and implementing technology for the Medicaid Enterprise. ▪ When necessary, recommend data management technologies and solutions applicable to the implementation of FX projects that align to MITA 3.0 and the Data Management Strategy. ▪ Collaborate with the SEAS Vendor on physical implementation of data store changes based on FX data model definition and updates.
FX Stakeholder Organizations	▪ Review and, as appropriate, may align technology solutions with FX data standards, systems, and processes per the Data Management Strategy to improve Medicaid program outcomes.

Exhibit 2-1: Roles and Responsibilities

SECTION 3 INFORMATION ARCHITECTURE FRAMEWORK

The FX future strategy is to establish an information architecture framework that will allow the Agency to use data exchange services, create a future-proof information foundation that can quickly adapt to evolving healthcare goals and processes, and position the Agency to readily adopt technology and solutions that do not yet exist. Connecting services and infrastructures and developing integration standards are an important initial step for advancing the level of MITA maturity and achieving the near-term goal of system modularity modernization.

The CMS released the Medicaid Program Final Rule: Mechanized Claims Processing and Information Retrieval Systems in December 2015. This final rule modifies regulations pertaining to the Code of Federal Regulations (CFR) Sections 42 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 its maturity in accordance with the MITA Framework in the Florida FX Procurement Strategy document. Interoperable Services, including key service interoperability elements, service model interoperability, and model architecture, is detailed in *T-5: Technical Architecture Documentation*, Section 5.4.6 (i.e., FX Hub > Standards & Plans > Category: Technology > Technical Architecture Documentation (T-5)).

The MITA Framework defines three interrelated architectures[1]:

- Business Architecture (BA) which describes the near- and long-term business operations of Medicaid
- Information Architecture (IA) which identifies the data needs of Medicaid
- Technical Architecture (TA) which comprises the technical elements used to specify new IT systems

Exhibit 3-1: MITA Framework Architectural Components shows a simplified view of the relationships the IA has with the BA and the TA. The BA describes the business processes along with data input, data output, and shared data required. The TA describes the technology enablers associated with various levels of maturity. Although the MITA Framework leaves the implementation of databases, data interchanges, and other physical components to the State MES, these are reflected in the Agency Data Management Strategy and closely interact with the three architectures as shown.

[1] Text colors correspond to Exhibit 3-1 below.

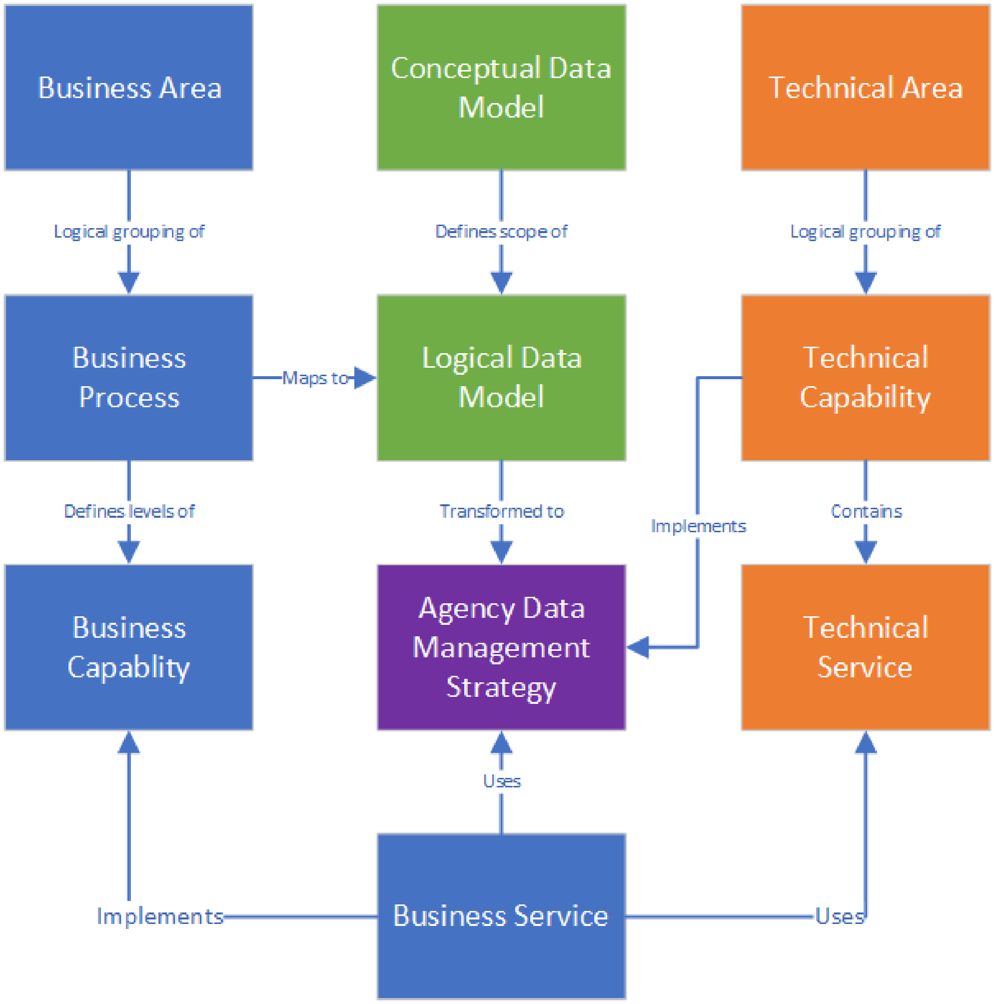

Exhibit 3-1: MITA Framework Architectural Components

The MITA Framework Information Architecture recommends data standards and identifies enabling technologies and interoperable designs for data exchange. To achieve this goal, FX leverages the MITA Framework to define a strategy for identifying, designing, and managing enterprise data across its information systems. Managing data across the enterprise requires strategic planning, coordination, and adherence to a shared set of principles to achieve the goals of interoperable systems and greater MITA maturity. Methods and activities specified in this deliverable incorporate the CMS MITA 3.0 framework for IA and expand on that standard by employing the broader government specific Government Information Factory (GIF) model developed by Bill Inmon.

GIF and MITA are complementary in that they operate at different levels of abstraction. The GIF is strategic in nature, encompassing the entire enterprise. MITA is focused more on tactical projects specific to Medicaid.

GIF provides an encompassing framework that:

- Enables an organization to identify the various information needs and interrelationships and the cornerstone elements that support the entire enterprise
- Addresses the unique needs of government agencies for data sharing, longevity, and security
- Remains technology neutral to not limit future implementations
- Provides for both near and long-term planning

The MITA IA components are tied to specific capabilities such as:

- Data Management Strategy (DMS) – Provides a structure that helps the development of information/data and promotes effective sharing across the Medicaid Enterprise to improve mission performance. The DMS is covered in more detail in the *T-1: Data Management Strategy* (i.e., FX Hub > Standards & Plans > Category: Technology > Data Management Strategy (T-1)).
- Conceptual Data Model (CDM) – Represents the overall conceptual structure of the data, providing a visual representation of the core data concepts needed to run an enterprise or business activity.
- Logical Data Model (LDM) – Represents the logical architecture of data that comprise or are shared within the Medicaid Enterprise. This model is comprehensive, detailed and technology neutral; and is subsequently transformed to a physical implementation model or models as appropriate.
- Data Standards (DS) – Discusses the available data standards and the benefits of using them. Standards for data definition, maintenance, and interoperability are contained in the *T-3: Data Standards* (i.e., FX Hub > Standards & Plans > Category: Technology > Data Standards (T-3)).
- Information Capability Matrix (ICM) – Defines the information capabilities used in a business process and informs the identification of technical capabilities.

3.1 Government Information Factory

In the wake of September 11, 2001, the unique needs for government agencies to integrate and share information triggered several initiatives including creation of the Government Information Factory. Unlike task focused solutions such as MITA, GIF was designed as a framework to support all the information needs of a government agency. The factors which shaped GIF, as outlined by Inmon, include:

- Exchange information with other government agencies
- Protect information held in confidence within the boundaries of the agency
- Provide online transaction processing between individuals and the agency
- Operate information processing within a finite budget

- Provide long-term archiving for certain kinds of records
- Provide public access through the Internet
- Protect agency processing from the intrusive access of data from the Internet
- Provide both proactive and reactive security through the agency's information systems
- Manage a large volume of data
- Monitor activity as it occurs
- Integrate data into a cohesive whole as it is collected from disparate sources
- Support the usage of sophisticated reporting and analytical tools
- Service many kinds of users across the agency's domain
- Provide accurate and timely information
- Provide data at a low enough level of detail that it can be reshaped to support all information needs
- Provide a definitive source of information in cases where the accuracy of information is questioned

Since the vision of FX reaches well beyond the MITA architecture, a more comprehensive framework is needed. **Exhibit 3-2: FX Information** (full size version in Attachment F) shows the various components and cornerstone elements of the FX Information Architecture Framework based on GIF. This framework is not built all at once; it is intended to guide the Agency in planning projects to implement incrementally and to situate those projects in perspective to their relationship to other projects and to the Agency as a whole.

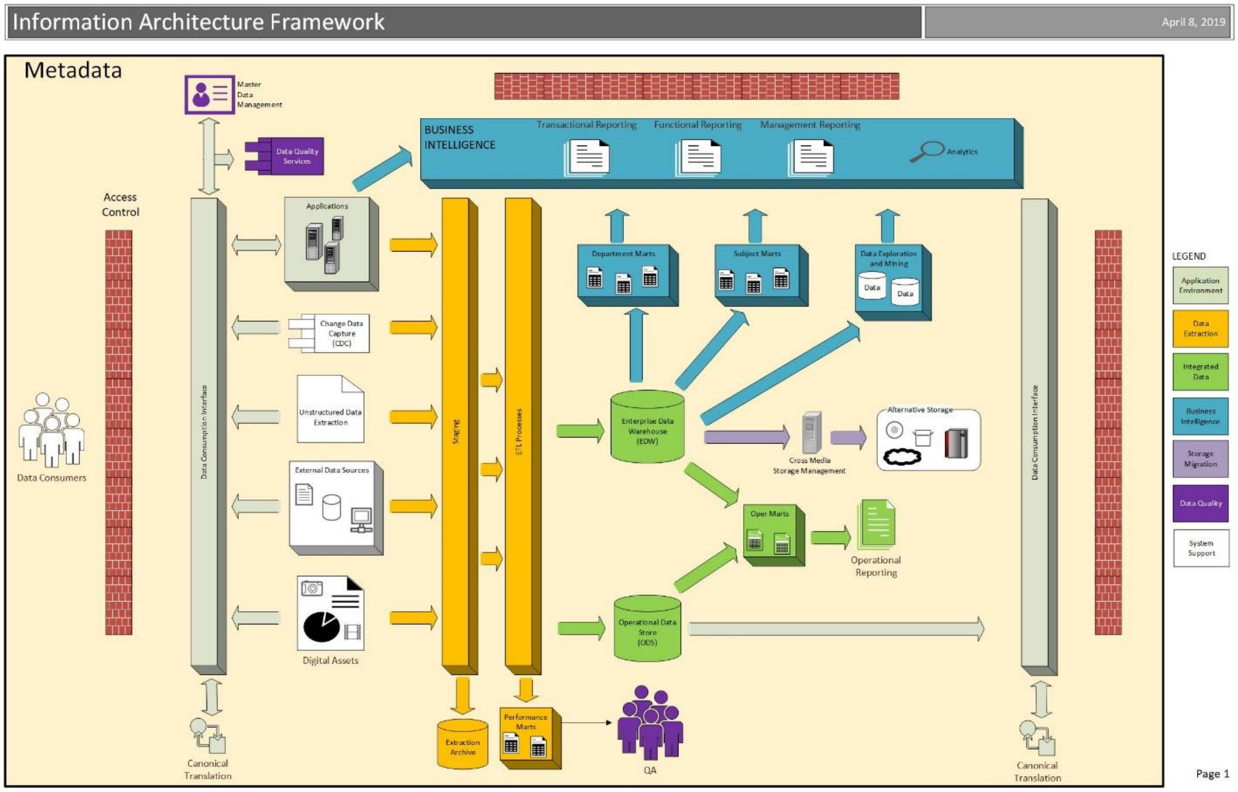

Exhibit 3-2: FX Information Architecture Framework

3.1.1 CORNERSTONE ELEMENTS

The legend identifies the cornerstone or *major* components of the framework.

- Application Environment – This is where the daily operations occur, and most data is created. Many interactions between the Agency and the user community are found in the applications.
- Data Extraction – The data extraction components serve two key purposes: 1) provide an auditable source of record for all data, and 2) remove volatility and variability from data values.
- Integrated Data – A critical issue in many organizations is data silos with inconsistent and sometimes contradictory data. The goal of data integration is to provide for the *single source of truth*.
- Business Intelligence – Provides for operational and analytical analysis necessary for informed decision-making.
- Storage Migration – Recognizes the cost/benefit of aging data and the storage media being used.

- Data Quality – Provides for master data, data cleansing, and monitoring of ongoing data quality.
- System Support – Covers people, technologies, and techniques that enable or enhance the data infrastructure.

Overarching the entire framework is the metadata about every component, process, and data element contained throughout the enterprise. This metadata is vital to understanding context, assessing and monitoring progress, and unifying the enterprise around a shared vision.

3.1.1.1 Analytic Architecture Framework

The purpose of defining an Analytic Architecture Framework is to provide direction to the Agency and its technology providers (e.g., FX module vendors) on the recommended approach to providing a sound, efficient, and cost-effective method for data acquisition and processing within the Agency's analytic environment.

Exhibit 3-3: The Technologist's Burden shows the many different technical aspects that need to be in place to meet the minimum analytic requirements of the business:

Exhibit 3-3: The Technologist's Burden

What is important to note is that ALL of the categories of components on the left must be in place to deliver business value, which is the goal of any analytic architecture. The analytic architecture provides a framework for understanding the scope and interrelationships of components involved in a data-centric organization and provides a basis for creating a roadmap of current and future data/analytic projects.

Exhibit 3-4: Reference Analytic Architecture depicts an Enterprise Data Warehouse (EDW)/analytic-centric approach for acquisition, loading, and access to EDW/Analytic data:

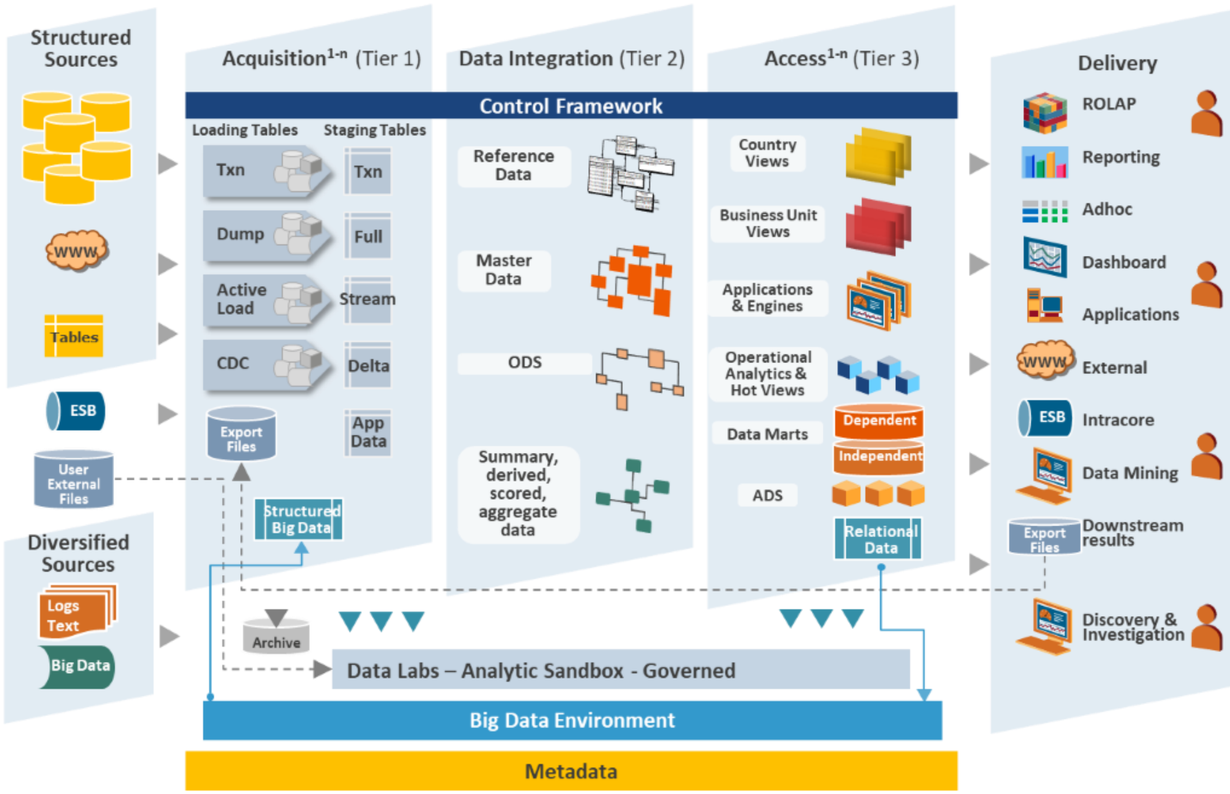

Exhibit 3-4: Reference Analytic Architecture

The Reference Analytic Architecture (RAA) is a high-level blueprint that shows the components and conceptual structure for building the Integrated Analytic Environment. The RAA is **EDW**, **or analytic**-**centric,** and only shows items such as the Enterprise Service Bus (ESB) and Big Data for context.

Please note that the above is a <u>reference architecture</u> and is not meant to imply any physical implementation other than a best-practice approach for creating and maintaining an analytic environment.

Data movement on this diagram generally moves from left to right. Although it is important to understand that the requirements begin on the far-right side, as in *What information needs to be delivered?* This, of course, is for known business questions. The flexibility in this architecture is to provide the Agency Enterprise with the ability <u>to answer any question of any data at any time!</u>

From left to right, source data is delivered through various data integration mechanisms, including batch file extracts, change data capture, streaming web data, and the ESB.

It is important to note the EDW/analytic structures at the center of the diagram consist of three layers or “tiers,” a Data Labs environment (to provide an area for users to engage in ‘what/if’ scenarios, combining their test data with production data), and a Big Data environment, typically a Hadoop distribution package.

- Tier One is the Acquisition or Staging Layer. Its purpose is to gather data from multiple source systems and serve as input for transformation processing. Historical data is kept here in case it is needed for future needs. The Archive is displayed under the Acquisition layer because data is kept in the acquisition area on a temporary basis and is archived after all data is moved into the Integrated Data Layer. In this case, *Archive* does not represent a system backup, which would be a much broader archive that includes the data and objects in all three tiers of the system. Data Quality and Master Data Management (MDM) processing occurs here. Records that fail these checks are suspended until a data steward can take remedial action. Note that the (1-n) designation in **Exhibit 3-4: Reference Analytic Architecture** indicates there are a number of different data sets kept in staging for a variety of business purposes.
- Tier Two is the Integrated Data or Core Layer. This is the physical instantiation of a logical data model. Its purpose is to store data in a single integrated data format, which is designed to promote cross-functional usage. Note that there are many different logical structures within the Core layer: Operational Data Store (ODS), Master data, reference data, etc. These are all components of the core model and are depicted as logical representations.
- Tier Three is the Access Layer. Its purpose is to transform the normalized and abstracted data found in the Integrated Data Layer into user-friendly structures. The proposed Analytical Data Store (ADS) sits here, although the determination if these reports will be virtually implemented or physically instantiated is not depicted. Note that views are a primary access pattern as access to the raw data in the core layer should not be directly accessed by the user community.
- Data Labs are the sand box area. Its purpose is to support rapid experimentation and evaluation of data without the formality of production tables. This is a governed area, which means that after exploration/discovery, the data and access will be removed.

The Metadata Definition & Lineage orange box at the bottom of the diagram reflects the importance of metadata to the data integration processes, especially for audit and lineage purposes. The Control Framework in **Exhibit 3-4: Reference Analytic Architecture** (across the top of the layers) is a generic depiction for some sort of job control. Think of A, B, C – Audit, Balance, and Control.

The **Big Data** box reflects the positioning of diverse, multi-structured data relative to the traditional relational EDW structures above it. Note that current approaches for unstructured data is to convert it to a more structured format. This makes it easier to integrate in order to gain the benefit from unstructured data.

All components fit together to create the Integrated Analytics Environment so that integrated data and big data can be made available to the business intelligence, data mining, reporting, and integration applications generically depicted on the right side of this diagram.
The whole point of architecture discipline is to make sure we are constructing designs that will address business priorities, initiatives, or pain (lack of needed insight and information). Architectural discipline is needed to ensure alignment between the business need, the existing capabilities of the organization, and the recommended solution/approach. This discipline also

yields benefits by way of lowered risk and lowered cost implementations, as well as providing consistency in delivery and resource ability and skill.

There are generic tool recommendations in the *T-3: Data Standards* documentation. The point here is to call out, from an architectural perspective, the use of the right tool for the right job. For example, use a Stored Procedure instead of an Informatica script to perform a function if it will meet the requirement.

3.1.1.2 PRINCIPLES

This section is based on the many years of experience in understanding the difference between what makes organizations successful with their analytics program and what makes other organizations not as successful.

Principles are generic rules, and in some cases, guidelines that form part of the framework within which an organization operates. Principles should rarely be changed, and when they are changed or deviated from the standard, this fact, together with the rationale for doing so, should be recorded.

The principles are the foundation for making decisions related to the architecture, thereby setting standards and resolving conflicts in situations where major decisions have potentially long-term consequences.

Exhibit 3-5: **Architectural Principles** lists the Architectural Principles and their rationale and benefits:

PRINCIPLE	STATEMENT	RATIONALE	BENEFITS
Simplicity	Where more than one option exists for the design then, if there is no reason to choose the more complex option, use the simpler option	Development, implementation, and maintenance of simpler systems is less expensive	Systems based on this principle should be smaller, easier to build, maintain, and govern throughout their life cycle
Abstraction	Used to distance a concept, component, or idea from a particular physical instance, thereby simplifying it and its explanation	Abstraction allows design to be independent of the physical instance and makes the design both easier to understand and more flexible to implement	Improved understanding between business and IT. Enables everyone involved, from business to developers, to properly understand how business requirements map to IT systems
Isolation	Separating functional components of the architecture into isolated units supports other principles such as simplicity and supportability	Isolating architectural components from each other makes them easier to build, easier to maintain, and easier to replicate for scalability	Overall productivity is enhanced by isolated components, allowing granular, parallel development. Changes within one component are possible without impacting other areas of the system. Isolated components are easier to replace or upgrade due to changes in technology or business requirements

Standards	Conforming to the principle of simplifying wherever viable, standards should be set and adhere to conformance	Defining standards for components reduces the initial development cost, including that of testing, which enables re-use and reduces potential confusion during operation and maintenance	Enables greater interoperability between systems. Overall costs of governance and maintenance are reduced
Scalability	The design should be such that increases in business demand may be met by replicating components within the IT system	An architecture which is capable of being scaled by replication of existing components, whether software or hardware	The ability to scale by replication will be a critical success factor for an architecture. The business can be confident that expansion of the use of IT resources can be met by simple, rapid duplication of existing components
Extensibility	Architectures should be designed such that they may be extended to meet future business requirements	Appropriate design decisions taken when originally creating an architecture can ease the future extension of the resultant solution into new areas of the organization	Business growth is facilitated by producing a system that is capable of extension without re-work, reducing cost and time to market
Supportability	All the architecture components should combine to produce an easily supportable architecture	This principle is implicit within several others, such as simplicity and standards -- it should be borne in mind at all times	Focus on supportability leads to the use of standard techniques, reducing both development cost/effort and risk of error
Integrity	The overall system should deliver integrity from source to target, with auditable processes	This leads to business confidence in the system's ability to deliver trustworthy, complete information	Integrity leads to reliability of answers to common, repeated business questions

Exhibit 3-5: Architectural Principles

Every architectural decision should be validated against these formal, mandated principles. Designs and variations are discussed and agreed to early-on.

3.1.1.3 Advocated Positions

Advocated Positions ensure the integrated data environment is built to a professional and flexible standard with long-term benefits at the lowest possible cost, and help balance between short and long-term architectural trade-offs. Advocated Positions are based on the deep, empirical knowledge that has evolved over time. Finally, they can (and do) change over time as new knowledge is gained.

Exhibit 3-6: Architecture Advocated Positions lists the Architecture Advocated Positions, their rationale, and benefits:

Position	Statement	Rationale	Benefits
Touch it/ Take it	When one or more attributes are required from a source file or table, all attributes in that file or table will be extracted.	Taking all the data from any given object is easier to implement and to change and significantly lowers cost	Greatly reduced costs when previously unspecified data (unspecified in the business requirements) is later required
Reversibility	The architecture should provide the capability for any data errors found in the core Physical Data Model (PDM) to	Sometimes business requirements mandate that data be reversed out of the	The design of the core PDM and supporting architecture components must be such that all

	be reversed out of the core PDM, with the core PDM being subsequently rebuilt correctly	system so it can be properly rebuilt	data may be identified by row, load process, and source system
Reusability	Wherever possible, common components should be used throughout the architecture to build infrastructure and applications	Combining the architecture principles of standards, simplification, and abstraction enables both physical and logical components, particularly the former, to be re-used rather than redesigned and redeveloped for each system or extension to a system	Development of common, or standard, assets reduces both development and support costs
Traceability	All core PDM production data should be traceable to its originating source system and thus fully auditable	As systems are increasingly required to produce financial and regulatory reports, it is often mandatory, by law, for such systems to be capable of producing full audit logs in business terms rather than IT terms, either regularly or on demand	This position effectively amounts to enabling full Basel II data and process reconciliation capability. For true auditing by currently accepted regulatory standards, the position requires audit at the row level or below
Collect Metadata	The architecture must provide for the gathering and retention of all available metadata, including business, operational, and technical information	Applying the advocated position of collecting and keeping metadata from the start provides the possibility of making use of it, whether or not a mechanism for doing so is part of the initial design at the deployment level	The position mandates that the initial metadata be collected, such that a fully-fledged, metadata-driven approach can be implemented in the future
Abstracted Core – Physical Data Model	The structure of the core physical data model should be abstracted from any specific business usage, the latter being handled in the access layer	This advocated position is a combination of several architecture principles, including particularly flexibility and extensibility. It is also the simplest architectural option	The Core PDM's primary function can be focused on storing data in its most natural, flexible state with minimal duplication. The core PDM will be available for all business queries as it will not be optimized for any one business area
Include Acquisition / Staging layer	All data architectures will include an acquisition/staging layer	This serves primarily as an audit function in that it must contain the source data as it was at the point of extraction, and thus it supports the traceability and auditability positions. Users will not be given access to this layer for production usage	This layer must be persistent in order to fulfill its audit and reconciliation functions. Conditioning of data and data profiling are permissible in this area since they are non-destructive. Transformation is not permitted as it is (potentially) destructive
No production reporting from the Data Lab/Sandbox	Essentially, the Data Lab is available for ad hoc data exploration, advanced analytics, and prototyping but not for production Business Intelligence (BI) reporting. It is	The Data Lab provides the facility, amongst other things, to evaluate data and its quality and usage opportunities, to develop prototype reports, or to perform one-time analysis	This position means that it is not possible to 'grab some data' from the acquisition layer, load it into the Data Lab, and then rapidly move this report into production

	governed, meaning it will be cleaned up after use.	for quick opportunity assessment and/or decision making	
Integrated logical and physical data models	The use of integrated, enterprise-wide data models for both logical data model (LDM) and core physical data model (cPDM) is essential to provide flexibility of both design, extension, and usage	Often, while the LDM is an integrated model and can flexibly and extensibility represent the entire business, the cPDM, produced by physical modeling of the LDM, is either less integrated or not integrated at all; this has the effect of 'throwing away' some of the benefits of an integrated LDM at the physical modeling stage	The LDM will not be used to produce a disjointed collection of data marts. This enables enterprise-wide analysis, i.e., the ability to ask any business question across the enterprise
Permanent archive of everything	All data, including metadata and data in the acquisition layer, must be archived or kept on-line. Many times, this is kept in the Staging (Tier 1) area, as data in Staging is basically the operational source data sans minor transformations	The ever-increasing dependence of operational business processes, as well as regulatory business reporting, on data environments means that any report or analysis may need to be reproduced at (virtually) any future date	Using this advocated position means that there will never be a situation where the required piece of data is that which the business does not have
Prioritize data access over data loading	All data stores, including the core PDM, should be designed to provide good access paths for query, rather than for data loading	Data usage is more important than data loading; it is, after all, the reason for building a data analytics environment in the first place. This applies as much to the core PDM as to the access layer components (which are naturally designed with data access in mind, by definition)	Less data will be redistributed at access time, taking up less time and resources. This provides good access paths for query and faster response times
Enforce referential integrity (RI)	The core PDM will use either hard or soft referential integrity	This position ensures that all data can be reported on to enhance the level of accuracy of answers to business questions at all times. It also avoids the potential complexity of outer joins for end users when accessing the core PDM	It is important to note that referential integrity may be either hard or soft; this is not advocating the use of hard, database-enforced RI in all situations The data warehousing environment can be trusted as having integrity and can be relied on as a trusted repository for regulatory and audit purposes
Full copy of source data objects in Staging	Wherever possible, the acquisition layers should contain a full copy of source data objects, as distinct from delta processing	Source system managed deltas can sometimes be unreliable and impose overheads on the source system. It is therefore desirable to have full copies of the data within each source table touched to populate the Staging layer	Source system select definition is made trivial by this advocated position, meaning that initial definition and subsequent absorption of source changes is reduced in cost and time. Where extensive history exists in the source, filtering will be necessary, or the source itself may contain a suitable sub-store of data

Single route to core Physical Data Model	Data should have only one route into the core PDM, meaning that a single file should have both one method and one periodicity for loading into Staging	This advocated position relates to the 'full dump' position, in that experience shows that some types of delta processes may miss changes that occur with logging off in a database source, resulting in integrity issues and the need for a synchronization process	Considerable examination should be made of the reliability of source systems when choosing to use delta load techniques
Load everything into core Physical Data Model	All data should be loaded into the core physical data model (PDM) and exceptions reported, or made available for further action	Increasingly, as Business Intelligence (BI) systems become more integrated with operational systems, they are used for regulatory reporting, need to be fully auditable, and are relied on for behavior-based analysis; for each of these uses, everything that has a place in the target model should be loaded	Full reconciliation to source systems is enabled by loading everything. Referential integrity issues may arise. Data quality issues will be identified and made available for analysis during loading

Exhibit 3-6: Architecture Advocated Positions

3.1.1.4 Structured Architectural Approach

The goal of a structured approach is to provide a holistic structure in which to build customized solutions for business intelligence and data management. We do this by providing a simple-to-understand set of design documentation usable by everyone involved in the business requirement and solution. We also look to provide a reference to enable the discussion of requirements and design using a standard, comprehensive structure, producing designs suited to the business problem in question.

Success starts with good architecture, and good architecture starts with the business. Unless the specifics of a project dictate otherwise, the **best practice** is to address the Business Architecture first, then the Information, Application, and System architectures. An easy way to remember this is *BIAS*, B-I-A-S, as in, we have a *BIAS* towards good architecture.

- **Business Architecture** provides the foundation for understanding customer needs in the context of the customer's current business state, including their business model, structure, mission, goals, and organizational processes.
  - What is the business model; where is the Business going; and how does it plan to get there?
- **Information Architecture** focuses on understanding what information is required to meet the identified business needs and includes an examination of sources, uses, storage, and organization of data.
- **Application Architecture** takes a look at each application – Extract, Transform, Load (ETL), Business Intelligence (BI), Advanced Analytics, and more – required to meet the business requirements.

- **Systems Architecture** evaluates the platforms, standards, and environments required to support the application, information, and ultimately the business architectures.
  - The Systems Architectures are the hardware, software, communications, facilities, and other related components, standards, and environments required to support the Application, Information, and ultimately the Business architectures.

The most important guideline here is to **always** start with the business and work through a set structure and process to produce "something which can be implemented." For many used to leading with technology, this approach may seem different and strange. It may take a little getting used to, possibly even a different mindset. However, experience dictates that in order to be successful with analytics, one must begin by isolating business needs before positioning a technical solution.

Exhibit 3-7: Architectural Ingredients for Success depicts the BIAS approach to architecture:

ARCHITECTURE CONTEXT	COMMENTS
B Business Architecture	▪ Defines organizational business model, structures, missions, goals, processes ▪ Business fundamentals are vital for organizational success
I Information Architecture	▪ Identifies data to support the Business View ▪ Includes calculations and rules ▪ Typically includes logical/physical data models ▪ Data is worked on by applications, used by business
A Application Architecture	▪ Application functions and needs ▪ Applications execute the functional side of the business architecture
S Systems Architecture	▪ The part IT cares about most ▪ Easiest to get **wrong** when we do not concentrate on the other aspects of architecture **first**!

Exhibit 3-7: Architectural Ingredients for Success

All architectural decisions must be driven by the business architecture and its requirements.

3.1.1.5 DESIGN PATTERNS

Design Patterns are re-usable approaches to solve commonly occurring problems. They are almost always logical, not physical in nature. In other words, they are the idea, rather than the implementation. Design patterns for an architecture are chosen based on requirements and context. Specifically, a pattern is a distinctive style, model, or form and is a model worthy of imitation.

The goal is to design a set of proven architectural options for meeting an array of requirements. Design patterns enable architects to have a starting point for evaluating architectural choices. Design patterns allow resources to avoid starting from scratch each time. They also provide a vehicle to capture expert knowledge on design trade-offs and implications. Additionally, they:

- Make expertise more widely available

- Promote design reuse
- Accelerate the design process
- Increase predictability
- Improve the quality of the design

3.1.1.6 Implementation Alternatives

Implementation Alternatives are the physical counterparts to design patterns. They are the embodiment of the implementation, rather than the idea. It is interesting to note that although these are often related to design patterns, they may exist independently.

Architecture is mainly about trade-offs. What one would gain from a certain architecture type must be balanced against what one might lose. If you think about it, you can “architect” something ten different ways, and they all may be technically correct. However, there are trade-offs that must be considered because each architectural solution may require different approaches to address the business need.

The following is an example to show the thought processes that should occur using a structured, formalized architectural approach to design. What is the business problem?
The problem is designing road junctions.

Our choices:

- Design patterns
  - Cross-roads
  - Cloverleaf intersection
  - Roundabout
  - Implementation alternatives for cross-roads design pattern

In this case, it is how to design an effective intersection for a given area considering the traffic patterns, flows, and what are the choices with which to decide. This is a well-known example that most can understand and put into context. Similar design issues exist with an integrated data environment… How does the data model need to be created so that it supports the business processes? How do folks access this data and how timely will it be? From the library of design patterns and implementation alternatives the Agency is building, what can we recommend in order to meet these needs?

Much of the trade-off in this architecture lies with data loading as opposed to data access. The difficulty in most EDW implementations lie with data access. Which is best to design for? **The answer is to design for data access**.

3.1.2 MITA Integration

The MITA Information Architecture comprises a subset of the FX Information Architecture Framework focused on specific components. First and foremost is the design of the data integration components and the Data Consumption Interface. Data integration is addressed through the data models representing the conceptual (semantic) view of the business and the logical (architectural) view of the enterprise data. Canonical Translation components of the Data Consumption Interface define the interchange data used to decouple the transmitted data from the underlying database design and the applications using the data.

3.2 Data Modeling

The FX Enterprise Data Models (Conceptual and Logical) support the entirety of the FX data in a manner that is independent of specific technologies or business processes. Properly designed, an LDM should accommodate changes to processes or technology platforms without needing to redefine the data components. Sub-models will be used to assist business process modeling by grouping the relevant entities along business area and functional lines. They will also be used to assist in physical design for selected data platforms.

Details about what is included in each model (entities, relationships, attributes, definitions, naming conventions, etc.) can be found in the *T-3: Data Standards* (i.e., FX Hub > Standards & Plans > Category: Technology > Data Standards (T-3)).

3.2.1 Modeling Strategy

Historically, the data designs behind the Medicaid Management Information System (MMIS) applications have been focused around optimizing application functionality in a monolithic system. The desired future state is to remove the dependency on applications for the definition of data. The modeling strategy is to capture a single, cohesive business view of data that accommodates both current operations and the near-term modularization of FX systems and enable new and expanded services for the enterprise.

An important distinction that needs to be understood is enterprise scope vs. enterprise wide. The goal of FX data modeling is to focus on the potential impact and utilization of data elements across the entire enterprise (wide) rather than create a complete model of every data element in the enterprise (scope). This allows for both incremental construction and faster realization of benefits.

3.2.2 Modeling Methodologies

Support of the FX data layer requires different modeling methodologies for the three key Enterprise Data Warehouse (EDW) environments: operational data, data warehouse, and data analytics. There will be different logical models for each based on a single enterprise conceptual data model. Physical data models are the responsibility of the FX EDW Vendor.

Operational data stored in the ODS is transactional data that is best modeled using an Entity-Relation approach. The preferred modeling method is Information Engineering (IE) due to its prominence in the industry as a standard for transactional systems.

The data warehouse is focused on supplying longitudinal data that has been integrated across multiple sources. At present, there doesn't appear to be driving factors (volume, velocity, variety) that would also necessitate a Big Data approach; however, the strategic approach is to use a modeling methodology that accommodates changes over time with minimal impact on existing data and little to no rework (e.g., conforming dimensions) needed. The preferred modeling approach for the data warehouse is Data Vault as it is already designed to accommodate high volume data loading, adding new data sources, and adapting Not Only SQL (NoSQL) / Big Data technologies with minimal impact to the existing environment.

Data analytics requires at least two different modeling methodologies: one for data marts and the other for data mining/exploration. Data marts will be modeled with star schema dimensional techniques to provide high performance querying and compatibility with common data visualization tools. Data mining/exploration areas use either a narrow-deep (few columns, many rows) or a broad-shallow data (many columns, few rows) set design that is best defined by the data analyst based on the problem space being analyzed.

3.2.3 Model Access

The *T-3: Data Standards* contains details on the location and access strategy for data models, data dictionary, and common vocabulary.

3.3 MITA Data Dimensions

The Medicaid Enterprise aligns to MITA's business areas as defined by CMS within the MITA 3.0 guidance. **Exhibit 3-8: MITA Data Dimensions** illustrates a detailed breakdown of the MITA business areas as defined by CMS. The headings in gold represent the ten MITA business areas and the headings in light green define the technical areas of integration and security respectively. This diagram illustrates the boundaries of the Medicaid Enterprise according to the MITA Initiative. Below each area is the breakdown of business or technical functions defined within the area.

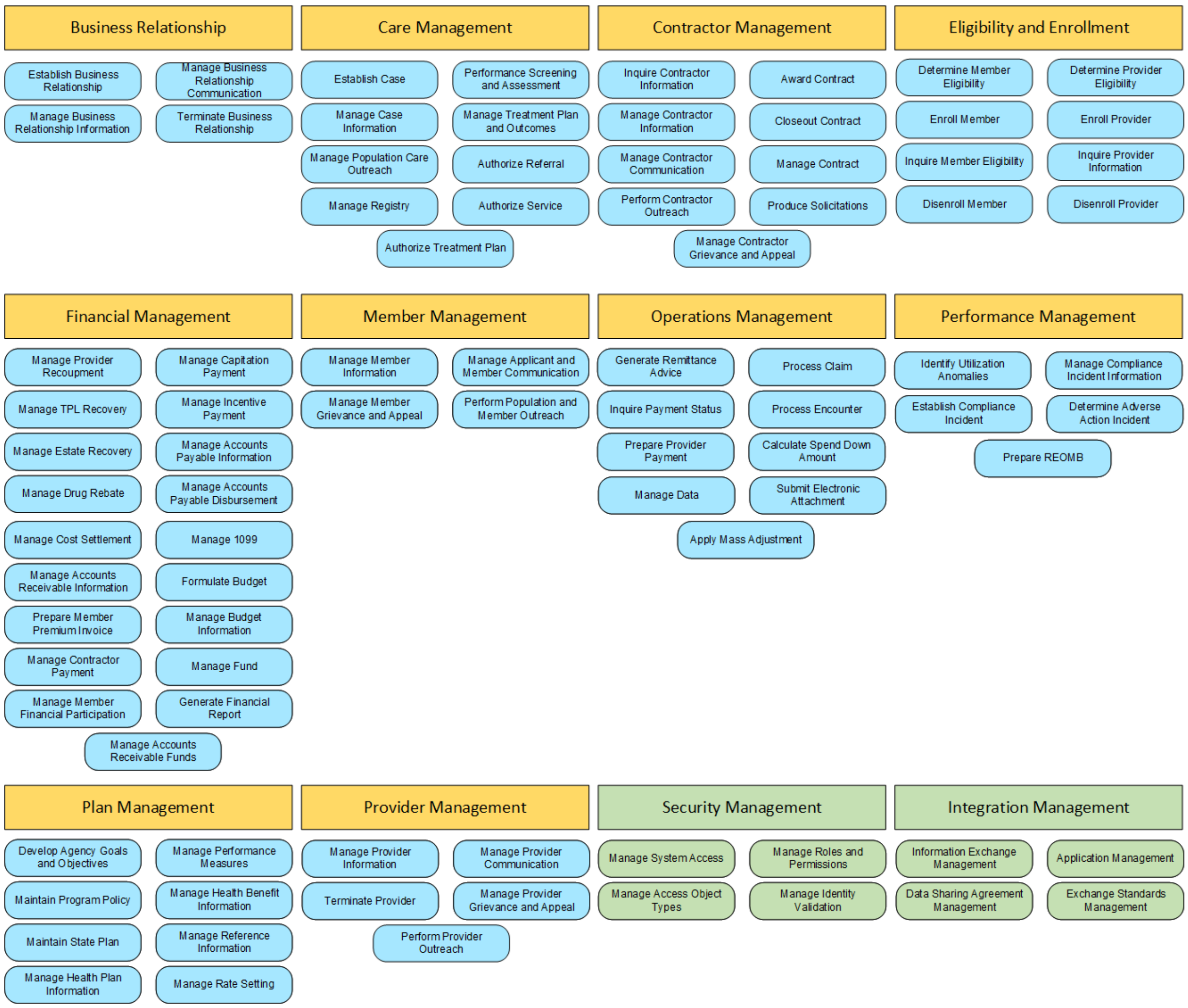

Exhibit 3-8: MITA Data Dimensions

Exhibit 3-9: Business Area Data Requirements provide a description of the major data categories required to support each business area.

Business Area	Data Requirements
Member Management	Member Management is a collection of information involved in communications between the Agency and the prospective or enrolled recipient and actions that it takes on behalf of the recipient. It also includes coordinating communications with both prospective and current recipients, outreach to current and potential recipients, and dealing with recipient grievance and appeals issues.
Provider Management	Provider Management is a collection of information involved in communications between the Agency and the prospective or enrolled provider and actions that the Agency takes on behalf of the provider. Business processes focus on terminating providers, communications with providers, managing provider grievances and appeals issues, and performing outreach services to providers.

Business Area	Data Requirements
Operations Management	Operations Management is a collection of information that manages claims and prepares buy-in premium payments. This business area uses a specific set of claims-related data and includes processing (i.e., editing, auditing, and pricing) a variety of claim forms including professional, dental, institutional, drug and encounters, and sending payment information to the provider.
Financial Management	Financial Management is a collection of information to support the payment of providers, health plans, other agencies, insurer buy-in premiums, and support the receipt of payments from other insurers, providers, and recipient buy-in premiums and financial participation.
Performance Management	Performance Management is a collection of information involved in the assessment of program compliance (e.g., auditing and tracking medical necessity and appropriateness of care, quality of care, recipient safety, fraud and abuse, erroneous payments, and administrative anomalies). This business area uses information about an individual provider or recipient (e.g., demographics, information about the case itself such as case manager ID, dates, actions, and status, and information about parties associated with the case) and uses this information to perform functions related to utilization and performance.
Business Relationship Management	Business Relationship Management is a collection of information that facilitates the coordination of standards of interoperability. This business area defines the exchange of information and Trading Partner Agreements (TPA) between the Agency and its partners, including collaboration among intrastate agencies, interstate agencies, and federal agencies. These agreements contain functionality for interoperability, establishment of inter-agency Service Level Agreements (SLA), identification of the types of information exchanged, and security and privacy requirements.
Care Management	Care Management information defines the needs of the individual recipient, plan of treatment, targeted outcomes, and the individual's health status. It also contains business processes that have a common purpose (e.g., identify recipients with special needs, assess needs, develop treatment plan, monitor and manage the plan, and report outcomes).
Plan Management	Plan Management information includes the strategic planning, policymaking, monitoring, and oversight business processes of the Agency. This business area is responsible for performance measures, reference information, and rate setting. The business processes include a wide range of planning, analysis, and decision-making activities. These activities include service needs and goals, healthcare outcome targets, quality assessment, performance and outcome analysis, and information management.
Contractor Management	The Contractor Management information accommodates a State Medicaid Agency (SMA) that has contracts or a variety of outsourced contracts. The Contractor Management business area has a common focus on Medicaid contractors (e.g., managed care, at-risk mental health or dental care, primary care physician), is responsible for contractor data store, and uses business processes that have a common purpose (e.g., fiscal agent, enrollment broker, Fraud Enforcement Agency, and third-party recovery).
Eligibility and Enrollment Management	The Eligibility and Enrollment Management business area is a collection of information involved in the activity for determination of eligibility and enrollment for new recipients, redetermination of existing recipients, enrolling new providers, and revalidation of existing providers. The Provider Enrollment business category and related business processes focus on recipient safety and fraud prevention through functions such as determining screening level (i.e., limited, moderate, or high) for provider verifications. These processes share a common set of provider-related data for determination of eligibility, enrollment, and inquiry to provide services. The Eligibility and Enrollment Management business area is responsible for the eligibility and enrollment information of the recipient data store and the provider data store.
Security Management	Security Management involves the management of authentication, roles and permissions, data sharing agreements, and standards required for compliance of the Medicaid systems.
Integration Management	Integration Management involves communications between the Medicaid Enterprise System and external trading partners such as partner agencies, modularized systems, external service providers, and outsourced application functions.

Exhibit 3-9: Business Area Data Requirements

3.4 USE OF DATA MODEL STANDARDS AND REFERENCE SOURCES

The FX data models will use appropriate standards provided by national standards organizations, State, and CMS guidelines where applicable. The FX data models will use applicable business language and common names preferred by national naming standards:

- National Information Exchange Model (NIEM) version 4, used in data interchange models
- Health Level 7 (HL7) Reference Information Model (RIM) datasets
- Federal Health Information Model (FHIM)

The NIEM version 4 provides the basis for modeling information exchanged between government organizations. NIEM is a canonical model that is the standard for exchange of information between government information sources. The federal government mandates use of NIEM for exchange of information between federal government organizations. NIEM is also widely adopted for modeling and exchange of information between other government data sources.

The use of NIEM based modeling and data vocabulary is a change from existing healthcare vocabulary that will require communication and organizational change management to secure adoption. The direction to use NIEM supports the direction for MITA higher maturity levels to increase integration and use of information across program, agency, and state boundaries. Likewise, the direction to use social determinants of care to increase coordination of health care will integrate new information sources, data types, and expand and change the vocabulary and perspective of the business. This evolution in vocabulary should be minor but will help particularly in communication with external organizations that already communicate with this vocabulary.

Exhibit 3-10: Data Modeling Use of Standards – NIEM, HL7 shows the subject areas defined in the NIEM canonical model.

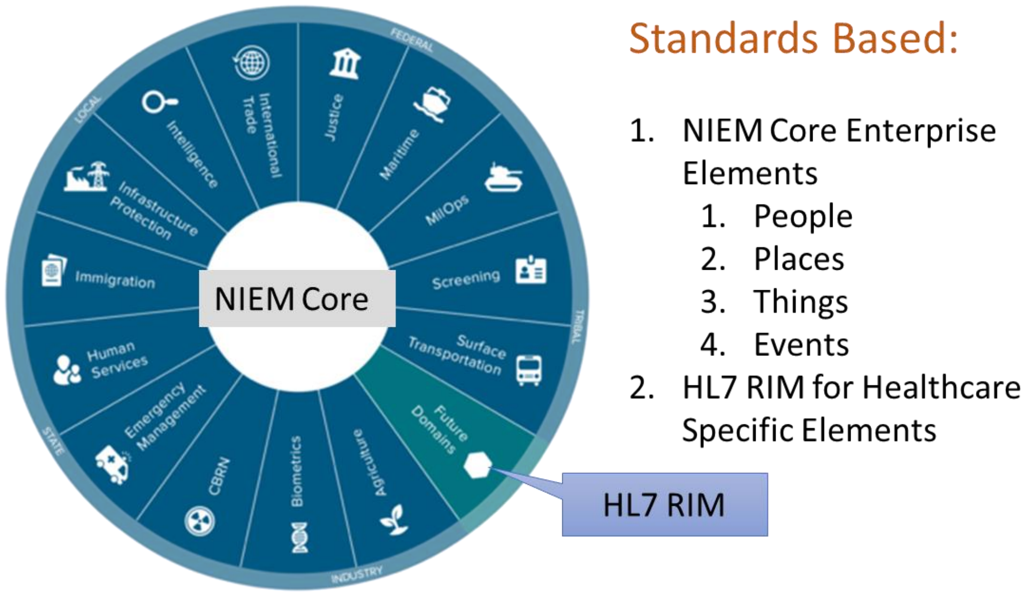

Exhibit 3-10: Data Modeling Use of Standards – NIEM, HL7

NIEM currently lacks data classes or elements defined for some of the healthcare specific data classes necessary for a SMA data interchange model, though a Health domain has been proposed. The preferred strategy is to make use of NIEM extensible namespaces rather than continue supporting multiple protocols and semantics. Other federal agencies are in the process of transforming their models to NIEM. For those not yet converted and other areas not presently supported by NIEM conventions, the data naming and formats will use industry standards such as those from HL7 rather than locally defined formats.

SECTION 4 FX CONCEPTUAL DATA MODEL

The FX CDM is a tool to bridge the knowledge gap between Medicaid subject matter experts, IT architects, and designers. The model depicts the major business information concepts and their relationships to each other, using business terminology for a business audience.

The strategy to develop the FX CDM is to iteratively model the conceptual data and relationships relevant to FX Project implementations. This approach begins with a high-level core model and incrementally adds conceptual model components for areas relevant to active FX projects. The current release of the CDM can be found in the FX Data Model Library (i.e., SEAS > Technical Domain > Data Management Depository > FX Data Model Library) and covers Medicaid, Legal Case Management, and Identity and Access Management.

CDM is not a solution model and is technology and application neutral. From a data perspective, the conceptual data model is a business model and not an architectural model for building databases. Business analysts use CDM to confirm and correct their understanding of the business. As CDM is a high-level model, attributes are rarely a component of CDMs unless necessary to clarify a business concept. CDM helps in establishing relationships between entities, though it will not provide the technical details about the relationship. CDM is independent of data storage technologies or database management systems.

4.1 FX Conceptual Data Model Approach

Although there are multiple ideas of what constitutes a conceptual data model, the FX approach is based on Information Engineering (IE). IE is the most frequently used approach by data modelers and practitioners. The IE definition of a conceptual data model includes the following characteristics:

- Enterprise-wide coverage of the business concepts
- Designed and developed primarily for a business audience
- Contains around 20-50 entities with no or very limited number of attributes
- Ideally, the entire model should fit on one page
- Contains relationships and may or may not include cardinality and optionality
- All entities have definitions
- Independent of databases, data storage, or other technologies; concepts may be implemented in either digital or non-digital methods

Exhibit 4-1: Sample Subset of a Conceptual Data Model shows the major components of a CDM. In the CDM:

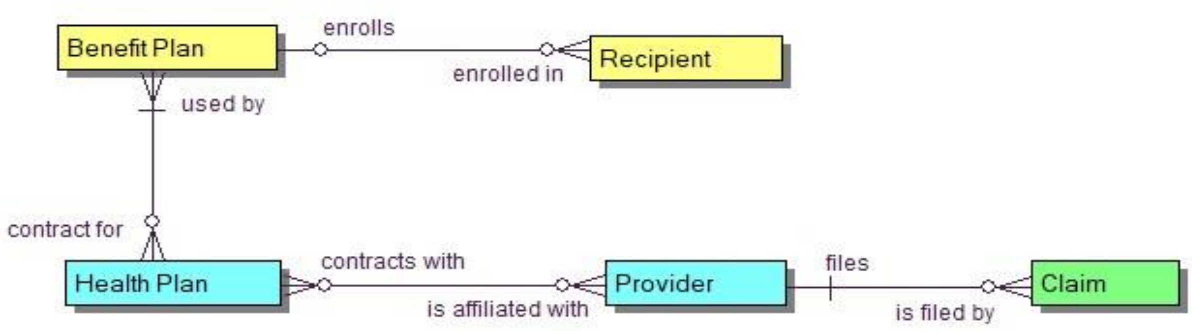

Exhibit 4-1: Sample Subset of a Conceptual Data Model

- Entities (boxes) represent real world things from a business perspective. This helps create a common vocabulary of core business concepts.
- Relationships (lines connecting entities) define the nature of business rules that affect how entities interact. Relationships are bidirectional and the verb phrase indicates the nature of the relationship from the perspective of each entity. For example, a Business Plan enrolls Recipients and a Recipient is enrolled in a Benefit Plan.
- Cardinality/Optionality (symbols on the relationship lines) express constraints on the relationship. Cardinality defines the maximum set size associated with an entity from the point of view of a single member of the other entity. A single Benefit Plan enrolls many (indicated by the crow's foot) Recipients. A single Recipient is enrolled in only one (indicated by no crow's foot) Benefit Plan.
- Optionality defines the minimum set size and is typically used to mean either zero (optional) or one (mandatory). A single Benefit Plan enrolls a minimum of zero (open circle at the crow's foot) Recipients (i.e., a Benefit Plan can exist without any Recipients being involved). A single Claim is filed by at least one (intersecting line) Provider (i.e., a Claim cannot be filed without a Provider identified).

A CDM helps identify key high-level business entities and the relationships existing between them. It also helps define the key issues of business problems and opportunities for the system.

4.2 FX Conceptual Data Model Content

Strategic Topic 4-1: FX Conceptual Data Model Content shows the strategic direction on content to include in the FX CDM and how much legacy data to reflect in the FX CDM. The implementation of this recommendation will be as an incremental release of the FX CDM.

Conceptual Data Model Content	Current	2018	Timeline 2020	2022	2025
Other agencies (beyond AHCA)		Exchanged Data	->		

CONCEPTUAL DATA MODEL CONTENT	Current	2018	TIMELINE 2020	2022	2025
Agency-wide all known current CDMs		Collection of As-Is existing CDMs			
Content from select CDMs					
Existing usage from EDW discovery					
To-Be usage following EDW discovery			To-Be	->	
Content based on rolling wave of FX projects		`	To-Be	->	
Analysis	The Agency sees value in having a current Agency-wide Conceptual Data Model (CDM) of all existing Agency data. As the Agency business areas drive changes in process and data usage, the Agency CDM would be updated to reflect the modeling needed to support the FX projects.				

Strategic Topic 4-1: FX Conceptual Data Model Content

4.3 FX CONCEPTUAL DATA MODEL RESPONSIBILITY

Strategic Topic 4-2: Conceptual Data Modeling Responsibility shows the strategic direction of who is responsible to perform conceptual data modeling for the FX CDM.

CONCEPTUAL DATA MODELING RESPONSIBILITY	Current	2018	TIMELINE 2020	2022	2025
SEAS Vendor		X	->		
EDW Vendor			Coordination with SEAS Vendor (2020)	->	
FX Project Owner			Coordination with SEAS Vendor (2020)	->	
TPA Vendor	FMMIS, DSS				
AHCA Agency Systems (e.g., IT, HQA, ….)	X				
Analysis	The SEAS Vendor is responsible for conceptual and logical data modeling. The FX Project Owners would coordinate with the SEAS Vendor to provide FX Project specific input. The EDW Vendor will be responsible for physical implementation and physical modeling. Coordination between the SEAS Vendor and EDW Vendor ensures a smooth transition.				

Strategic Topic 4-2: Conceptual Data Modeling Responsibility

While the SEAS Vendor is responsible for developing and maintaining the conceptual and logical data models, the EDW Vendor will also coordinate closely with the SEAS Vendor in the physical implementation of data models and data services.

SECTION 5 LOGICAL DATA MODELS

The LDM is a fully attributed model that serves as a foundation to enable the reengineering of FX business processes. It is derived from the conceptual data model and defines the structure of information independent of any data base management system (DBMS), technology, data storage, business process, or organizational constraints. Using an enterprise wide, shared LDM helps the Agency achieve integration and interoperability across modules.

The LDM provides:

- Focus on what data comprises the organization, and not on what data is needed by the processes
- Facilitation of business-focused data analysis
- Aid in understanding enterprise-wide business rules and business data usage, and uncovering existing data defects from a 360-degree (360º) view of a business
- A basis for performing data integration
- Improved data quality

Per MITA guidance, the LDM also provides the basis for generation and implementation of physical model(s), which describe how data will be structured to meet performance objectives in a specific physical implementation.

5.1 FX Logical Data Model Characteristics

The FX LDM represents the abstract structure of a domain of information. A single, enterprise-wide logical model, however, would quickly become unwieldy. Consequently, the FX LDM will use sub models to aid in identifying the information components of specific subject areas and providing working models for developers building applications. All sub models will be fully synchronized to the enterprise-wide model; the concept is primarily a display mechanism to ease use.

Common characteristics of an LDM include:

- Typically contains 100-500 entities, depending on the scope of the model
- Designed for both business and technical audiences
- Relationships contain cardinality and optionality constraints
- Attributes will have logical data types with length and precision defined
- Attributes will have optionality defined
- Entities, attributes, and relationships will have definitions
- Is independent of any specific database or data storage model
- Contains significant metadata beyond just the diagram itself (the iceberg model)

- Is derived from the semantics but not necessarily the architecture of the CDM

A major factor in the design of the FX LDM is a direct result of decoupling the data from specific applications. The design of the decoupled data makes use of universal data model components (e.g., Party, Role, Communications) and is abstracted enough to future-proof the model from changes in functionality or business processes.

Decoupling also has an impact on the type of modeling being done. The approach shown in MITA documentation is based on UML (Universal Modeling Language), an object modeling technique used by application designers. The industry standard methodology for data modeling, however, is Entity Relationship Diagramming (ERD). ERD is a static modeling method depicting data independent of processing. UML, when used to depict data models, is focused primarily on the physical layer rather than the conceptual or logical models. UML contains details necessary to implement behavior in both code and databases but is missing or overly complex in defining certain logical core data concepts. Although the two methods are complementary, they are not readily interchangeable as each differs in their central paradigm. FX data modeling is focused on creating a singular source of enterprise-wide data that is housed separately and isolated from the current business processing logic to achieve modularity of applications, prevention of data lock-in by a vendor, improved data quality through *single source of truth*, and faster decision-making through reduced latency in data availability. Consequently, FX will be using ERD rather than UML for data modeling.

5.2 FX Logical Data Model Content

Strategic Topic 5-1: FX Logical Data Model Content shows the strategic direction on content to include in the FX LDM and how much legacy data to reflect in the FX LDM. The implementation of this recommendation will be implemented as an incremental release of the FX LDM.

LOGICAL DATA MODEL CONTENT	Current	2018	TIMELINE 2020	2022	2025
Other agencies (beyond AHCA)		Exchanged Data	->		
Agency-wide all known current LDMs			Collection of As-Is existing LDMs		
Content from Select LDMs					
Existing usage from EDW discovery					
To-Be usage following EDW discovery			To-Be	->	
Content based on rolling wave of FX projects			To-Be	->	
Analysis	The Agency sees value in archiving current LDMs of all existing Agency data. As the Agency business areas drive changes in process and data usage, the existing LDMs will be instrumental in defining data lineage and data transition plans needed to support the FX projects.				

Strategic Topic 5-1: FX Logical Data Model Content

5.3 FX LOGICAL DATA MODEL RESPONSIBILITY

Strategic Topic 5-2: Logical Data Modeling Responsibility shows the strategic direction on who is responsible to perform logical data modeling for the FX LDM.

LOGICAL DATA MODELING RESPONSIBILITY	Current	2018	TIMELINE 2020	2022	2025
SEAS Vendor		X	->		
EDW Vendor			Coordination with SEAS Vendor	->	
FX Project Owner			Coordination with SEAS Vendor		
TPA Vendor	FMMIS, DSS				
AHCA Agency Systems (e.g., IT, HQA, …)	X				
Analysis	The SEAS Vendor is responsible for conceptual and logical data modeling. The FX Project Owners coordinate with the SEAS Vendor to provide FX Project specific input. The EDW vendor will be responsible for physical implementation and physical modeling so coordination between SEAS and EDW vendor ensures a smooth transition.				

Strategic Topic 5-2: Logical Data Modeling Responsibility

While the SEAS Vendor is responsible for developing and maintaining the conceptual and logical data models, the EDW Vendor will also coordinate closely with the SEAS Vendor in the physical implementation of data models and data services.

SECTION 6 INFORMATION CAPABILTY MATRIX

6.1 INFORMATION CAPABILITY MATRIX OVERVIEW

The Information Capability Matrix (ICM) is an important component of the State Self-Assessment (SS-A). The ICM describes the IA component at a specific level of MITA maturity for each business area. The MITA Maturity Model defines five (5) levels of maturity showing how the Medicaid Enterprise will evolve. The high-level capability descriptions are as follows:

- **Level 1** Capabilities – Are predominantly manually intensive, IA components that do not take advantage of current industry standards.
- **Level 2** Capabilities – Are a mix of manually intensive components and electronic transactions or automated functionality internal to the Agency.
- **Level 3** Capabilities – Adoption of a governance process, a CDM, an LDM, enterprise modeling, the MITA Framework, industry standards, and other nationally recognized standards for intrastate exchange of information. Partners include one or more state agencies.
- **Level 4** Capabilities – Include interoperability amongst all appropriate state agencies, regional partners, regional Health Insurance Exchange (HIX), regional Health Information Exchange (HIE), and other external regional healthcare stakeholders.
- **Level 5** Capabilities – Include interoperability amongst all appropriate state agencies, regional partners, federal agencies, national Health Insurance Exchange (HIX), national Health Information Exchange (HIE), and other national external healthcare stakeholders.

Exhibit 6-1: Information Capability Matrix Template shows the detailed questions assessed for each MITA business area for the following Information Architecture areas:

- Data Management Strategy (DMS)
- Conceptual Data Model (CDM)
- Logical Data Model (LDM)
- Data Standards (DS)

The ICM describes the components as-is and to-be capability maturity at different points in advancement toward maturity (e.g., movement from local code sets (Level 1) to national code sets within the Agency (Level 2) to national code sets within all state agencies (Level 3) to the use of clinical data (Level 4) to the use of nationally adopted standards (Level 5)).

In performing the 2018 SS-A update, the Agency provided input for all ICM template questions for each MITA Business Area.

Exhibit 6-1: Information Capability Matrix Template shows the MITA ICM template.

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 and Medicaid Services (CMS) and other national agencies and groups to promote sharing of Medicaid resources.
Does business area have common data architecture?	No standards for data architecture development.	Implementation of internal policy and procedures to promote data documentation, development, and management where the SMA defines data entities, attributes, data models, and relationships sufficiently to convey the overall meaning and use of Medicaid data and information.	Adoption of intrastate metadata repository where the SMA defines the data entities, attributes, data models, and relationships sufficiently to convey the overall meaning and use of Medicaid data and information.	Adoption of a regional metadata repository where the SMA defines the data entities, attributes, data models, and relationships sufficiently to convey the overall meaning and use of Medicaid data and information.	Adoption of a national centralized metadata repository where the SMA defines the data entities, attributes, data models, and relationships sufficiently to convey the overall meaning and use of Medicaid data and information.

Business Area Title					
	Level 1	**Level 2**	**Level 3**	**Level 4**	**Level 5**
Does each business area use Enterprise Modeling?	No enterprise modeling exists.	Implementation of Medicaid internal policy and procedures to promote enterprise modeling.	Adoption of intrastate enterprise modeling to promote standardized data across data source systems and third-party resources to decrease resource expenditure and increase enterprise knowledge.	Adoption of regional enterprise modeling to promote standardized data across data source systems and third-party resources to decrease resource expenditure and increase enterprise knowledge.	Adoption of national enterprise modeling to promote standardized data across data source systems and third-party resources to decrease resource expenditure and increase enterprise knowledge.
Does business area use data sharing architectures?	No sharing of data.	Development of Medicaid centralized data and information-exchange formats.	Adoption of statewide standard data definitions, data semantics, and harmonization strategies.	Adoption of regional mechanisms used for data sharing (i.e., data hubs, repositories, and registries).	Adoption of national mechanisms used for data sharing (i.e., data hubs, repositories, and registries).

Conceptual Data Model (CDM)					
Does business area have CDMs?	No CDM developed.	Adoption of diagrams or spreadsheets that depict the business area high-level data and general relationships within the Agency.	Adoption of a CDM that depicts the business area high-level data and general relationships for intrastate exchange.	Adoption of a CDM that depicts the business area high-level data and general relationships with regional exchange including clinical information.	Adoption of a CDM that depicts the business area high-level data and general relationships with national exchanges.
Logical Data Model (LDM)					
Does business area have LDMs?	No LDM developed.	Identification of data entities and attributes relationships, data standards, and code sets within the Agency.	LDM identifies the data entities, attributes, relationships, standards, and code sets for intrastate exchange.	LDM identifies data entities, attributes, relationships, standards, and code sets for regional exchange including clinical information.	LDM identifies data entities, attributes, relationships, standards, and code sets for national exchange.
Data Standards (DS)					
Does business area use structure and vocabulary data standards to support current and emerging health data standards?	The Agency uses non-standard structure and vocabulary data standards.	SMA implements internal structure and vocabulary data standards used for performance monitoring, management reporting, and analysis. SMA implements state-specific and Health Insurance Portability and Accountability Act of 1996 (HIPAA) data standards.	SMA standardizes structure and vocabulary data for automated electronic intrastate interchanges and interoperability. SMA implements MITA Framework, industry standards, and other nationally recognized standards for intrastate exchange of information.	SMA standardizes data for automated electronic regional interchanges and interoperability. SMA implements the MITA Framework, industry standards, and other nationally recognized standards for clinical and interstate exchange of information.	SMA standardizes data for automated electronic national interchanges and interoperability. SMA implements the MITA Framework, industry standards, and other nationally recognized standards for national exchange of information.

Exhibit 6-1: Information Capability Matrix Template

6.2 2018 ICM Documentation

The 2018 ICM assessment performed interviews of key Agency representatives from different business areas and areas of expertise. The results of the 2018 assessment are in an ICM document available as part of this deliverable (in Attachment A: ICM 2018) Update.

The document sorts the first column alphabetically by MITA Business Area.

SECTION 7 MAINTENANCE APPROACH

The information architecture documentation is intended to guide the FX Project teams in the design and construction of data stores that support the FX conceptual and logical data model structures. Rather than store the various artifacts in a configuration management system, the information architecture itself will be maintained in this document and its attachments, and the data models (CDM, LDM, and historical archived models) will be maintained in the data modeling tool and replicated in the FX Data Model Library that is part of the FX Data Management Depository.

The FX Conceptual and Logical Data Models will evolve over the course of the FX life span. The SEAS Vendor maintains these data models. The SEAS Technical Lead works with Agency Business Units and other FX stakeholder organizations identified as owners of one or more business area data models for socializing the information architecture. The SEAS Vendor will also communicate and promulgate these conceptual and logical data models across future system design projects, as appropriate, using data modeling tools.

Agency leadership is also expected to provide direction on data modeling topics and participate in decision-making of issues escalated through the FX governance process.

7.1 Inputs and Outputs

Inputs and outputs define the reference material used to maintain the FX data models.

Exhibit 7-1: Input and Output Documents lists the documents and their purpose.

Document	Purpose
MITA 3.0 Part II Chapter 3 Conceptual Data Model	▪ INPUT: The Conceptual Data Model (CDM) is a blueprint or conceptual plan for building an information system's IA. The CDM serves as a tool that enables the reengineering of business processes and enterprise strategies.
MITA 3.0 Part II Chapter 4 Logical Data Model	▪ INPUT: The Logical Data Model (LDM) provides the mechanism for ensuring the completeness of the business model and serves as a tool that enables the reengineering of Medicaid business processes.
MITA 3.0 Part II Chapter 6 Information Capability Matrix	▪ INPUT: The Information Capability Matrix (ICM) is one of the principal building blocks of the MITA Framework. Business and Technical Services use information enabled by the IA capabilities.
IT MITA Fundamentals	▪ INPUT: Fundamentals of MITA 3.0 CMS Perspective provided alignment of business areas to anticipated models.
Information Engineering	▪ INPUT: Information Engineering is the most widely used methodology for conceptual and logical data modeling.

DOCUMENT	PURPOSE
NIEM v4	▪ INPUT: Core data naming elements and definitions for enterprise data across the Agency. ▪ NIEM version 4.0 is current; began as the result of a collaborative effort by the U.S. Department of Justice and Department of Homeland Security to produce a set of common, well-defined data elements to be used for data exchange development and harmonization.
HL7 Reference Information Model	▪ INPUT: A healthcare specific model of information with enterprise-data applicability. ▪ The HL7 Reference Information Model (RIM) is the cornerstone of the HL7 Version 3 development process. An object model created as part of the Version 3 methodology, the RIM is a large, pictorial representation of the HL7 clinical data (domains) and identifies the life cycle that a message or groups of related messages will carry. It is a shared model between all domains and is the model from which all domains create their messages. The RIM is American National Standards Institute (ANSI) approved.
SEAS Deliverable T-1: Data Management Strategy	▪ INPUT: Provides the strategic direction for the management of data within FX.
SEAS Deliverable T-3: Data Standards	▪ OUTPUT: Contributes to the detailed data standards and dictionary to support the development of the physical data model.

Exhibit 7-1: Input and Output Documents

7.2 Document Maintenance

The Medicaid Enterprise is continually evolving along with new legislation and technology. Even as the State Medicaid Enterprise evolves, increased functionality, tighter performance standards, and anticipated health outcomes will continue to change business operations and the technology used to conduct business. The Agency's plan is to achieve higher levels of capability in various timeframes. The MITA Framework ICM encourages growth and transformation by illustrating the benefits of improving FX operations and provides tools to help the Agency achieve that transformation.

7.2.1 New Model Designs

As FX continues its evolution through the addition of system modules and improved business processes, new models may be developed to support that implementation. The FX data dimensions, as defined above, serves as a guide for the development of new models.

7.2.2 Updates to Existing Components

Requestors seeking to update or incorporate new content in the IA and/or FX data models will create a Project Change Request (PCR). The PCR provides visibility to the change history.

The *T3: Data Standards, Section 3.1.15* describes the Change Management Process. This process should be used to make changes to IA and/or FX data models.

7.3 Data Model Maintenance Process

Currently, the Agency has multiple stores of conceptual and logical data model information (FMMIS, AHCA IT, others). The data models for each area have different levels of detail, discipline, update frequency, and use different processes. The strategy is to maintain the FX data models in a central location to control the process of making changes. Data models are currently stored in the FXPR at SEAS > Technical Domain > Data Reference Content > Data Management Depository > FX Data Model Library. Storage of data models in this location is considered an interim solution. If FX acquires a fully integrated suite of data management tools with a common repository, it will no longer be needed.

7.4 Information Capability Matrix Development

This section defines the development and maintenance process used in the 2018 ICM and the strategy for future ICM development.

7.4.1 Traditional ICM Development and Maintenance

The development of the 2018 ICM used the standard process of SS-A development and updates used by most states. The standard SS-A development and update process is a periodic assessment process performed annually. The process pulls together groups of business and technology resources to answer standard questions that define the maturity of a business area and capability. Overall maturity levels are defined for specific questions with an overall category level defined as the lowest assessment value for each category. For example, the Conceptual Data Model category addresses questions concerning business area data management, common data architecture, enterprise modeling, data sharing and model existence. Each question defines and specifies a maturity level based on the answer to that question (such as no sharing of data is Level 1). The maturity level for the entire category is defined as the lowest maturity level found across those questions, so if the answer to data sharing is none (Level 1), then the entire category is Level 1.

The traditional SS-A process has several drawbacks including:

- Significant business and technology resources participation requirements
- To-be state projected independent of project approval and project funding
- SS-A cost

7.4.2 Future ICM Maintenance Strategy

The FX strategy is to shift to near real-time, event-based maintenance and update of the ICM. The approach is to embed the assessment of ICM changes in maturity into the FX Portfolio Management Process. As the strategic project portfolio management process evaluates potential FX projects for approval and implementation, the business and technology resources that are defining and validating business benefit and impact can quickly identify and document the effects of the FX Project on MITA ICM maturity. When each proposed FX Project has an associated change in MITA ICM maturity, it will be possible to automatically generate the projected to-be maturity level for SS-A reporting required by CMS as of any date in the future based on FX Project scheduled implementation dates.

APPENDIX A – SUPPORTING DOCUMENTATION

The following attachments are stored in the FXPR to serve as supporting documentation for the Information Architecture Documentation deliverable. (i.e., FX Hub > Standards & Plans > Category: Technology > Information Architecture (T-2))

ATTACHMENT A - ICM 2018 UPDATE

ATTACHMENT B - FX CONCEPTUAL MODEL DATA DICTIONARY

ATTACHMENT C - FX CONCEPTUAL MODEL MEDICAID DIAGRAM

ATTACHMENT D - FX CONCEPTUAL MODEL CASE MANAGEMENT DIAGRAM

ATTACHMENT E - FX CONCEPTUAL MODEL IDENTITY AND ACCESS MANAGEMENT DIAGRAM

ATTACHMENT F - INFORMATION ARCHITECTURE FRAMEWORK