AHCA Florida Health Care Connections (FX)

T-3: Data Standards

Version: 300

Date: July 29, 2022

Author: The SEAS Vendor

Submitted To: AHCA FX Program Administration Team

Revision History

DATE	VERSION	DESCRIPTION	AUTHOR
4/30/2018	001	T-3: Data Standards Development Draft Version (Entry)	Jessica Rein Steve Quante Paul Moore
5/24/2018	002	T-3: Data Standards Revisions to address deliverable review comments	Jessica Rein Steve Quante Paul Moore
6/5/2018	003	T-3: Data Standards Revisions to address deliverable review comments	Jessica Rein Steve Quante Paul Moore
6/7/2018	100	T-3: Data Standards approved and baselined version	Sean Gibbs
5/30/2019	101	T-3: Data Standards Revisions to refine and expand topic areas for annual deliverable refresh	Barry McConnell
7/30/2019	102	T-3: Data Standards annual deliverable refresh response to review comments	Barry McConnell
7/31/2019	200	T-3: Data Standards approved annual deliverable refresh	Carol Williams
12/17/2019	201	T-3: Data Standards quarterly refresh	Barry McConnell Steve Ruszkowski
1/9/2020	202	T-3: Data Standards Revisions to address Agency deliverable review comments	Barry McConnell Steve Ruszkowski Joanna Redding
1/24/2020	203	T-3: Data Standards Revisions to address Agency deliverable review comments	Barry McConnell Steve Ruszkowski Joanna Redding
1/27/2020	225	T-3: Data Standards approved quarterly deliverable refresh	Eric Steinkuehler
5/31/2022	226	T-3: Data Standards draft refresh updates including DET #86 and #434: ▪ Updated Section 1 boilerplate language to reflect current approved standard language ▪ Updated Exhibit 2-1 to include EDW Vendor per DET #434 ▪ Defined Metadata Management Tool ▪ Updated Strategic Topics 3-3, 3-6, 3-8, and 3-9 ▪ Updated Project Change Request (PCR) to reflect Module Change Request (MCR) ▪ Updated Change Control Board (CCB) to reflect Module Control Board (MCB) ▪ Updated Exhibit 3-4 ▪ Updated Sections: › 3.1.9 › 3.1.11.3 (DET #86) › 3.1.12	Amelia McHenry Dave Schiller

DATE	VERSION	DESCRIPTION	AUTHOR
7/29/2022	300	T-3: Data Standards approved final	Carol Williams

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
4/30/2018	Sean Gibbs	QA Submission Review
5/25/2018	Sean Gibbs	QA Resubmission Review
6/5/2018	Sean Gibbs	QA Resubmission Review
5/30/2019	Carol Williams	QA Resubmission Review
7/30/2019	Paul Moore	Domain Resubmission Review
12/16/2019	Eric Steinkuehler	QA review
5/21/2022	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

The purpose of the *T-3: Data Standards* is to develop and establish the FX Data Standards to facilitate the interoperability of systems and effective data sharing. The FX Data Standards align with the *MITA 3.0 Part II Information Architecture – Chapter 5 Data Standards* while accounting for unique Agency requirements. The FX Data Standards are the product of current state discovery, stakeholder input, strategic analysis, program strategy, and direction about techniques and priorities to support overall improvement of Medicaid program outcomes.

The FX Data Standards document may contain paths to later versions of documents and diagrams, referenced within the following sections and used within this document, which reside in the FX Projects Repository.

The primary audiences for FX Data Standards are interested stakeholders including Agency for Health Care Administration (AHCA) technology leadership, executives, CMS, existing and potential FX Project Owners, FX vendors, Medicaid stakeholder agencies (e.g., other Florida Agencies), consumers of Medicaid data, and other state Medicaid programs.

1.3 SCOPE STATEMENT

This iteration of the deliverable focuses on the technologies, processes, and tools needed to implement the *MITA 3.0 Part II Information Architecture – Chapter 5 Data Standards* with emphasis on the data standards related to the foundational capabilities of Integration Services and Integration Platform (IS/IP), Enterprise Data Warehouse (EDW), and modular capability implementation. This document is not a detailed implementation manual but provides the context, aligned with MITA, required for planning purposes.

The FX Data Standards communicate strategy and direction for the following topics:

- Data Management
- Data Standards Coordination Process
- Data Standards Assessment Process
- Data Strategy, Architecture, and Standards Update Process

1.4 GOALS AND OBJECTIVES

The Goals and Objectives of this document are:

Goal 1 – Establish the MITA compliant FX Data Standards.

- Objective 1 – Define and document each of the core data standards areas for the Agency.
- Objective 2 – Provide key strategic data standards guidance and reference for future FX procurements as part of the Agency's modular implementation approach.

Goal 2 – Establish processes to assess data standards and maintain the data standards repository.

- Objective 1 – Provide the structure and taxonomy for the FX Data Standards Repository.
- Objective 2 – Provide guidance on the systematic process for assessments and updates to FX Data Standards.

1.5 REFERENCED DOCUMENTS

Documents referenced to support the development of this document include the following:

- Documentation in the Agency's FMMIS documentation system (iTrace)
- FX Projects Repository (FXPR)
- SEAS S-3: Enterprise Systems Strategic Plan
- SEAS T-1: Data Management Strategy

- SEAS T-6: Technology Standards
- MITA 3.0 Part II Information Architecture – Chapter 5 Data Standards
- MITA 3.0 Front Matter FM6 – Introduction to the MITA Framework
- MITA Information Series – Medicaid Overview
- Florida Planning, Accounting, and Ledger Management (PALM) Data Management Plan

1.6 STRATEGIC TOPIC INVENTORY

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

The Strategic Topic Inventory has many features to communicate a spectrum of strategic direction options considered across time for a specific topic. The timing of the strategic topics aligns to when the guidance is expected to begin to take effect or begin to be procured or implemented (e.g., IS/IP, EDW, Modules, etc.). Many factors outside the scope of this document will determine when any FX capability would be implemented. A summary chart can dynamically display the strategic direction for a specific topic across the time spectrum from current state direction to direction for future years. The Strategic Topic Inventory includes a summary analysis that describes the context and considerations that influenced the defined strategy for each specific topic.

Extracted topic specific summary charts from the Strategic Topic Inventory tool are included throughout this document to communicate recommended strategy and direction for many of the important data standards decisions that are important for FX stakeholders to understand.

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

Exhibit 1-1: Strategic Topic Inventory Item Sample shows a screenshot of a sample 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

The SEAS Vendor developed and maintains this Microsoft Excel-based tool that resides as a document in the FXPR.

SECTION 2 ROLES AND RESPONSIBILITIES

Exhibit 2-1: Roles and Responsibilities identifies the roles and responsibilities for the primary stakeholders involved with this deliverable.

ROLE	RESPONSIBILITY
SEAS Vendor Data Architect	▪ Identifies the evolving data standards necessary to improve the Medicaid Enterprise ▪ Reviews and proposes new FX Data Standards, updates, and the retirement of obsolete Data Standards to the FX Technology Standards Committee (FXTSC) for Governance Committee approval ▪ Identifies the risks of adopting new FX Data Standards
EDW Vendor	▪ Populates and maintains the FX Data Dictionary ▪ Maintains the FX Data Dictionary Change Request process ▪ Maintains the physical data models ▪ Evaluates and processes FX Data Dictionary change requests ▪ Creates and updates conceptual data models ▪ Creates and updates logical data models
FX Technology Standards Committee	▪ Creates specific rules to help identify new data standards ▪ Reviews, recommends for Governance Committee approval ▪ Proposes new, updated, and retired FX Data Standards
Data Governance Workgroup	▪ Reviews and approves/denies proposed new and updated FX Data Dictionary entries ▪ Oversees Data Dictionary enhancements ▪ Extracts FX Data Standards and FX Data Dictionary artifacts for use in FX Project procurements
Agency Data Stewards	▪ Documents the origin and sources of authority on data elements ▪ Defines and documents the business glossary data elements ▪ Responsible for utilizing data governance processes to ensure fitness of data elements and metadata
Agency Data Owners	▪ Assumes ultimate accountability for data assets ▪ Establishes guidelines and protocols governing proliferation, security, access, retention, archival, and disposal of data elements ▪ Ensures compliance to regulations, policies, and standards ▪ Defines data controls to manage risk
FX External Organizations	▪ Reviews and as appropriate may align technology solutions with FX Data Standards to improve Medicaid enterprise outcomes ▪ Contributes recommendations for FX Data Standards to improve integration, interoperability, consistency, and coordination
FX Project Owners	▪ Communicates using the technology vocabulary in the Technology Standards Reference Guide (TSRG) in proposing, discussing, and implementing technology for the Medicaid Enterprise ▪ Identifies and understands FX Data Standards applicable to implementation of projects using vendor provided technology or software
Integration Services / Integration Platform Vendor	▪ Creates and updates physical data models for the Integration Platform ▪ Coordinates extensions to canonical model in cooperation with EDW Vendor
EDW Vendor	▪ Creates and updates physical data models for EDW and module implementations ▪ Coordinates extensions to canonical model in cooperation with EDW Vendor

Exhibit 2-1: Roles and Responsibilities

SECTION 3 DATA CONCEPTS

Per the MITA Information Series – Medicaid Overview, CMS MITA guidance reinforces the importance of proper data definition to FX:

> "A universal data dictionary and standard definitions of common data elements will help MMIS transcend platforms. Using 'best of breed' systems for special purposes require that these individual systems be compatible with the MMIS' data and architecture standards, so they can communicate directly with each other and the resulting processed data will be meaningful when merged into operational data stores."

3.1 DATA MANAGEMENT

Data management includes a set of processes and tools used to organize and define the structure and attributes of data that FX creates, uses, and manages. These include the elemental components of FX Data Models and the FX Data Dictionary along with data lineage, transformation rules, security and compliance classifications, data domains, constraints and limitations, and other metadata components. The FX Data Repository is a collection that houses the discrete data concepts and relationships that underlie the Agency business and the conceptual, logical, and physical data models. This section explains the organizing principles and strategic technology direction to manage and communicate the entries that provide definition and understanding to users and processors of FX information.

Exhibit 3-1: FX Data Management Components highlights the major components and component relationships that organize and provide structure for FX Data Management.

At the highest level, data categories and data subject areas provide the structure to organize data components. Defined industry, governmental, and subject area models often provide reusable components for baseline models. The FX Data Dictionary is the repository for data element definitions. FX data models use data element definitions to organize data into sets that may be stored, exchanged, or processed together. The FX Data Standards provide guidance and rules for the data definitions in the FX Data Dictionary and the data models in the FX Data Repository.

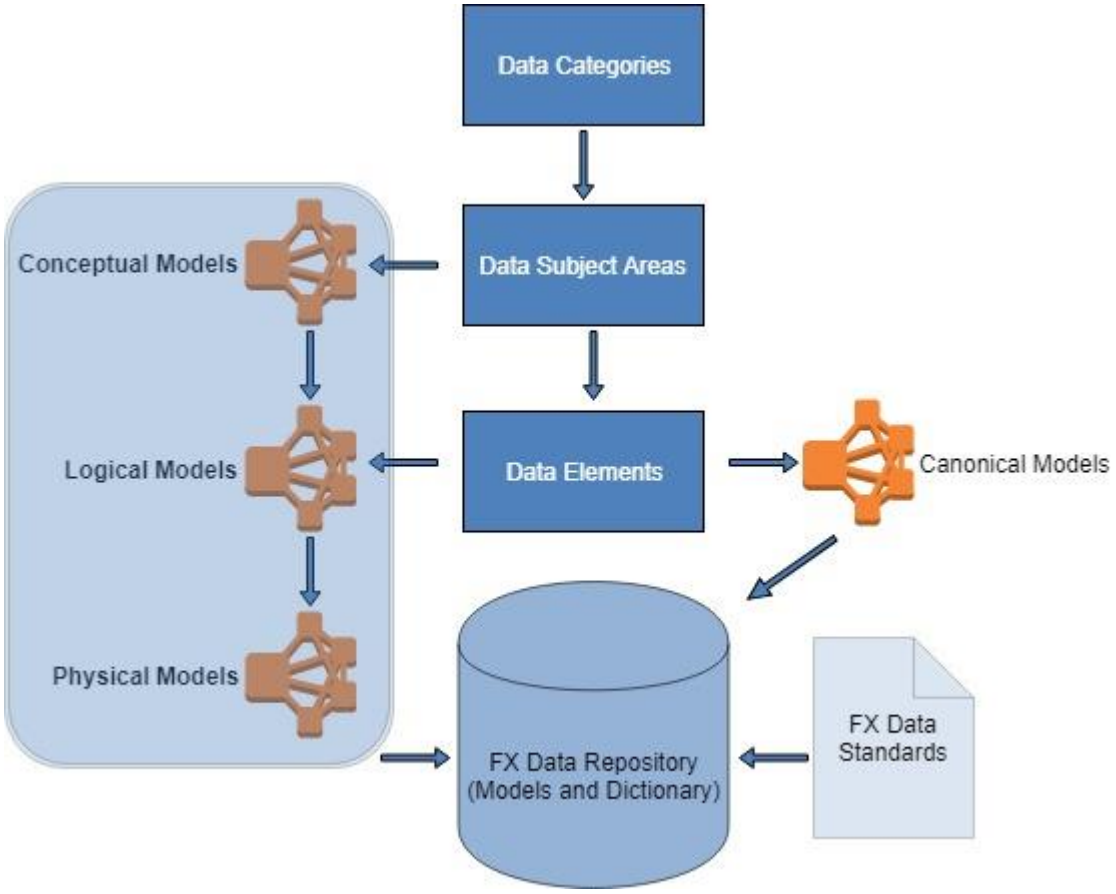

Exhibit 3-1: FX Data Management Components

3.1.1 DATA CATEGORIES

Data Categories are the highest-level organization of data that provides broad groupings aligned with data usage profile and data location. The main categories of data used in the FX are:

- Operational Data
- Information Exchange Data
- Analytic Data
- Rule and Policy Data
- Experience Data
- Program and Project Management Data

3.1.2 DATA SUBJECT AREAS

Data Subject Area is a lower-level grouping that provides structure to data in a data category. Subject areas group data into related topics aligned around a business use or major object class. The first 10 subject areas listed below are MITA Business Areas as described in MITA 3.0 Front Matter FM6 – Introduction to the MITA Framework. **Exhibit 3-2: Sample Data Subject Areas** displays example subject areas.

SUBJECT AREA	DESCRIPTION
Member Management	A collection of information gathered in communications between the Agency and the prospective or enrolled recipient and actions that the Agency takes on behalf of the recipient.
Provider Management	A collection of information gathered in communications between the Agency and the prospective or enrolled provider and actions that the Agency takes on behalf of the provider.
Operations Management	A collection of information about the management of claims and the preparation of premium payments.
Financial Management	A collection of information to support the payment of providers, health plans, other agencies, insurers, and premiums. Supports the receipt of payments from other insurers, providers, and recipient premiums and financial participation. Supports the creation of federal financial reporting (e.g., CMS 64).
Performance Management	A collection of information about 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).
Business Relationship Management	A collection of information about the coordination of business operations between organizations that deliver services.
Care Management	A collection of information that defines the needs of the individual recipient, plan of treatment, targeted outcomes, and the individual's health status.
Plan Management	A collection of information about strategic planning, policymaking, monitoring, and oversight business processes of the Agency.
Contract Management	A collection of information about Agency health plan contracts or outsourced contracts.
Eligibility and Enrollment Management	A collection of information about the activity for determination of eligibility and enrollment for new applicants, redetermination of existing recipients, enrolling new providers, and revalidation of existing providers.
Security Management	A collection of information about the management of authentication, roles and permissions, data sharing agreements, and standards required for compliance of FX systems.
Integration Management	A collection of information about communications between the FX and external trading partners such as external organizations, modularized systems, external service providers and outsourced application functions.

Exhibit 3-2: Sample Data Subject Areas

3.1.3 DATA ELEMENTS

A data element is an atomic unit of data that has a precise definition or semantics and conveys meaningful information to its user. Any unit of data defined for processing is a data element,

but the term is generally used in conjunction with a specific context (e.g., communications data element, provider data element, etc.). Data elements contained in business data models are referred to as attributes and are associated with entities (business concepts) and relationships (business rules).

Note: The reuse of the same data element definition in multiple data models for multiple data categories is likely and acceptable.

3.1.4 CONCEPTUAL DATA MODELS

A conceptual data model is a summary-level model that describes the entire enterprise. The purpose is to organize, scope, and define business concepts, rules, and the relationships between enterprise business areas. Although it may contain some attributes to provide context or clarification, the conceptual model does not provide complete details about the data elements involved.

3.1.5 LOGICAL DATA MODELS

A logical data model is a fully attributed model that describes data requirements from the business point of view and is technology neutral. The purpose is to develop a complete map of business rules and business concepts and provide a comprehensive description of the business. Logical sub-models may be organized around specific subject areas and integrated into an enterprise-level model.

Although logical models are typically associated with entity relational design and relational databases, they are not implementation specific. They are useful for both dimensional and Not Only SQL (NoSQL) (so-called *schema-less*) databases. The strategy for FX is that all data will be logically modeled and transformed into the appropriate physical models.

The misperception that there is no need to model with NoSQL approaches is because of timing. Data Lake, Hadoop, and similar Big Data approaches focus on loading data first and deriving the meaning later. While this solves the problem of capturing high volume, high velocity data, it does nothing to optimize using the data. Significant effort is expended in deriving a model (key-value, graph, document, etc.) and extracting the data into a usable form for analysis. However, when the data has already been defined in a logical model, the physical transformation can focus on the iterative, query-optimized modeling appropriate to NoSQL databases (Cloudera, MongoDB, Neo4j, etc.).

3.1.6 PHYSICAL DATA MODELS

A physical data model is a representation of a data design proposed or implemented in a specific database management system. Derived from the logical model, a physical model includes the definition of database management system-specific structures and implementation-specific attributes such as domain, length, constraints, and performance-oriented attributes such as storage location, indexes, and partitioning. It may also contain

security and access control attributes. The physical model can also extend the logical model to include specialized data elements for system support and performance enhancements.

3.1.7 CANONICAL MODELS

Canonical models are a specialized type of data model that presents data elements in their simplest possible physical form without regard to specific application/database implementations. They are a design pattern based on a common set of definitions, values, and rules used in the exchange of information across systems. Canonical models define data in motion as opposed to data at rest like the previous models. They support system and database integration processes and data exchanges between different systems. A key characteristic is that these models are independent of the technology used by either the applications or services that use them.

The Agency sends data to and receives data from multiple systems, hence the strategy for utilizing a canonical information exchange model. The benefits of a canonical model over ad-hoc or application-specific strategies include:

- Structured data definitions that have standardized naming, standardized structures, and are better organized and documented
- Reduction in data translations and maintenance efforts
- Improved system interoperability

A widely used canonical model is the National Information Exchange Model (NIEM) that is the United States federal standard for government information exchange and used across all levels of government and private industry. This model includes a core model and multiple subject area models with data definitions, relationships, and predefined information exchange structures.

3.1.7.1 NIEM

NIEM originated as an XML-based, extensible information exchange model. In addition, XML NIEM supports JavaScript Object Notation (JSON) and Unified Modeling Language (UML) modeling. The NIEM model provides a set of terms, definitions, and formats for many business concepts and rules for how these business concepts align to each other. The NIEM model is independent from how information is stored in individual systems. NIEM canonical models will be defined and form the basis of the information exchange across modules and subject areas in the FX. For example, information exchanges utilizing person or organization identifying data will use data element definitions based on the NIEM canonical model.

Exhibit 3-3: NIEM Reference Model Subject Areas shows the subject areas in the NIEM model. The NIEM core area includes cross-cutting concepts like person and organization and may be used as a linkage between specific subject areas of information. A significant benefit of these core elements is that as the Medicaid Enterprise builds maturity in providing and using a 360-degree view of person and organization information, FX could integrate data across the

NIEM subject areas that provide social determinant of health data which, in turn, can improve coordination of care.

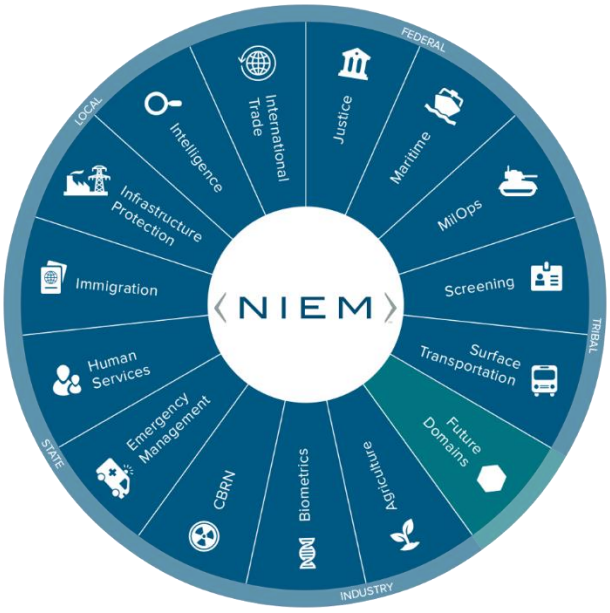

Exhibit 3-3: NIEM Reference Model Subject Areas

Strategic Topic 3-1: Preferred Canonical Model Strategy defines the Agency position on using a canonical model.

PREFERRED CANONICAL MODEL	TIMELINE				
Alternatives	**Current**	**2018**	**2020**	**2022**	**2024**
XML schema defined but not used consistently across all source and destination systems and services	FMMIS	FMMIS			
Use an existing data model as the canonical data model					
Use a canonical data model			X	->	

PREFERRED CANONICAL MODEL	TIMELINE
	Currently, the Agency does not use a canonical model pattern for the exchange of information across systems. Common industry data formats are used (e.g., X12) which are mapped to XML schemas, but those schemas are not consistently reused within FMMIS.
Analysis	The future state strategy is to use a canonical model pattern for the exchange of information across systems. This pattern will be implemented as part of the Enterprise Service Bus (ESB) which will be able to translate standard transaction types (e.g., X12, Fast Healthcare Interoperability Resources (FHIR), HL7, and National Council for Prescription Drug Programs (NCPDP) D.0) to support the exchange of data between modules and systems.

Strategic Topic 3-1: Preferred Canonical Model Strategy

Strategic Topic 3-2: Canonical Model Extensions describes how external organizations will extend the FX canonical data models.

CANONICAL MODEL EXTENSIONS	TIMELINE				
Alternatives	**Current**	**2018**	**2020**	**2022**	**2024**
External organizations will define their own model and not extend from the base model	X	X	X		
External organizations will extend from the inherited base model in a common (non-Agency) namespace					
External organizations will extend from the inherited base model in their own namespaces				X	->
	Currently, the Agency does not use a canonical model pattern for the exchange of information across systems. Common industry data formats are used (e.g., X12) which are mapped to XML schemas, but those schemas are not consistently reused within FMMIS.				
Analysis	The future state strategy is to use a canonical model pattern for the exchange of information across systems. Extensions to the base model will be accommodated through inheritance. Inherited models (e.g., for external organizations) will reside in their own namespace. A namespace is an organization technique to group objects of similar type. In this example, each external organization will have its own namespace for their agency specific model extensions.				

Strategic Topic 3-2: Canonical Model Extensions

3.1.8 FX DATA STANDARDS

FX Data Standards provide the rules and guidance for FX Data Management. The FX Data Standards defined in this deliverable focus on those rules that are relevant to data element/model definition and FX Data Dictionary use.

The sections that follow describe the FX technology strategy and direction specific to FX Data Management.

3.1.9 DATA ARTIFACTS MANAGEMENT

FX projects will create data artifacts using multiple tools. To simplify locating these artifacts, the FXPR site page titled Data Management Depository (i.e., SEAS > Technical Domain > Data Management Depository) provides a central location for accessing the artifacts directly or to obtain extracts created from the originating tools. The term depository is used to highlight that this is a place where artifacts are deposited for consumption. IBM's Information Governance Catalog (IGC) is the metadata repository tool used for the EDW solution. IGC will be the first place technical and business users will go to ascertain aspects of the data model, data dictionary, business glossary, data lineage, and other metadata components to help inform the end user Agency community of the data flow, transformations, business rules, and other associated metadata items. As IGC will be the master repository of current metadata information, it does not store history natively. As such, the Agency has determined that the history of documents, data models, mappings, etc., will reside on a SharePoint directory for reference.

3.1.10 FX DATA REPOSITORY AND DATA DICTIONARY

The FX Data Repository provides the business vocabulary and technical specifications to understand context and use of FX data. The FX Data Repository is a centralized collection of business terms, data elements, models, and metadata (e.g., constraints, data lineage, transformations, value domains, security classification, etc.) that is used by various tools (data modeling, data steward glossaries, metadata managers, etc.) to store and control access to these artifacts. The repository is a centralized content store for data management artifacts such as business glossary, metadata, data lineage, transformations, and models (process and data). It typically supports a suite of tools. The tools often have overlapping functions and storage may be local or in the repository to allow their sale as stand-alone components (e.g., metadata managers and modeling tools both contain a data dictionary concept though with different contents). In an integrated suite, they share a common data dictionary in the repository. As an interim solution, the depository in the FXPR provides access to a centralized store of artifacts. If FX acquires a fully integrated suite of data management tools with a common repository, it will no longer be needed.

The Data Dictionary defines the elements used in the data models and is maintained by the data modeling tool, which uses the FX Data Repository as its storage mechanism. The Data Dictionary is an effective and concise way to describe the elements used in data stores and data exchanges.

The FX Data Repository provides:

- Authoritative definition of the meaning and characteristics of data elements
- Centralized model management and integration across subject areas
- Consistent vocabulary that improves communications and accelerates FX development
- Improved data quality and consistency
- Improved documentation and control of FX data and models
- Enhanced ability for data consumers and processors to understand data that exists in data stores and is used in data exchanges throughout the enterprise

The IGC will serve as the metadata and data repository once the EDW solution is active.

3.1.11 FX Data Repository Strategy

The following section explains the strategy for definition and use of FX data elements by:

- Articulating the overall strategy for documenting data for FX including attributes and metadata for each data element.
- Defining expected vendor use of data definitions to design and implement FX modules and data management components.

Agency Data Stewards and Data Owners play a critical role in representing the business interests related to the creation and maintenance of data dictionary content. The *T-1: Data Management Strategy,* Section 4.1 *Data Management and Data Stewardship* provide additional details on Data Steward and Data Owner roles and responsibilities.

Key elements of the FX Data Repository Strategy include:

- Use of a centralized FX Data Repository – The FX Data Repository is the single source for models and data definitions supporting FX related projects within the Agency. The expectation is the FX Data Repository eventually supports multi-Agency definitions of all data relevant to the FX.
- Consolidation of existing Agency data dictionaries into the FX Data Repository – Existing data definitions in system specific repositories will be migrated to the FX Data Repository.
- Centralized administration – Centralized administration of the FX Data Repository with models, data element definitions, and data lineage encourage consistency and promotes reuse. The SEAS Vendor will monitor the volume, types and process to analyze change requests.
- Distributed data modeling – Modeling of conceptual, logical, and physical implementations will be a distributed function performed by multiple FX Project Owners and the maintainers of source systems. The SEAS Vendor will be

responsible for conceptual and logical models, the EDW and IS/IP Vendors will be responsible for physical models and canonical models. Module vendors will be consumers of the FX data models rather than creators. The Agency will be accountable for data design rather than the module vendors.

- Retention or conversion of data definition history – The FX Data Dictionary will convert definition history that can be migrated from existing data definition history sources or will retain the history in the native data definition source for reference.
- Integration between the Metadata Management Tool (Information Governance Catalog (IGC)) and FX Data Repository – The metadata management tools and data modeling tools both include data dictionary capabilities. The FX Data Repository uses the data dictionary capability in the data modeling tool because it is more relevant to data modelers and should be more cost effective to use. The FX Metadata Management Tool will have the ability to integrate with the data modeling tool to reference the Agency's data dictionary information. All EDW users will have access to both the data dictionary and metadata to provide a holistic view of Agency data definitions. The *T-1: Data Management Strategy,* Section 8.3 *Enterprise Data Management Tools* discusses Enterprise Metadata Management Tools.
- Using context-based canonical models as the basis for exchange element definitions – FX will use data definitions from industry, federal, CMS, and state canonical models for data definitions of information used in data exchanges.
- Linking element synonyms that have different names in other models or industry standards – The FX Data Repository will capture element synonyms.

3.1.12 Modeling Approaches

Depending on the category of data being modeled, special factors must be considered. For example, operational data is processed as a single unit (a claim, a provider, a recipient, etc.) whereas analytical data accesses potentially millions of data records at once. These differences require wholly different approaches to performance tuning, storage volume, data retention, and level of detail. Sections 3.1.12.1 through 3.1.12.5 below describe the kinds of data contained within each data category and the approaches to take when modeling.

The SEAS Vendor will perform conceptual and logical data modeling for all data. Physical data modeling for the FX Integration Platform will be performed by the IS/IP Vendor. The EDW Vendor will perform physical data modeling of the data elements contained in the EDW and module implementations. The IS/IP and EDW Vendors will coordinate the canonical model and extensions needed.

3.1.12.1 Factors in Operational Data Modeling

Operational data modeling (more commonly known as OLTP or Online Transactional Processing) is characterized by individual transaction records used in applications and systems (e.g., claims, recipient, and provider data).

The FX approach to operational data is based on key core ideas including:

- All application data across modules going into the single Operational Data Store (ODS)
- Decoupling data from proprietary systems and databases via services
- Centralized data storage and control

3.1.12.2 FACTORS IN INFORMATION EXCHANGE DATA MODELING

Information exchange data modeling (also known as canonical modeling) is modeling the information passed between systems or services independent of application or database technologies. The data in an information exchange is transitory and relies on a structure such as NIEM to provide guidance for proper formatting and interpretation. The elements used in the information exchange data model include:

- Agreed upon terms, definitions, relationships, and formats
- Rules and methodology for using the model
- Extensibility for local customizations

3.1.12.3 FACTORS IN ANALYTICAL DATA MODELING

Analytic data modeling is the modeling of information to support decision-making, policy evaluation, and longitudinal studies. This information is typically held in the enterprise data warehouse, data marts, and reporting data stores. The most common modeling method for analytical data is dimensional modeling, though recent trends in data mining/data science defer modeling until after data collection in unstructured or semi-structured components like data lakes.

Analytic data modeling follows one of three core methodologies:

- Inmon (old style – near 3rd Normal Form), - The Inmon approach is based on the idea that the data warehouse is an integrated, non-volatile, subject-oriented database that serves as a single source of data for all analytical functions and downstream data sets such as data marts.
- Data Vault (endorsed by Inmon to replace his old style) - The Data Vault modifications provided a better modeling method than using relational 3rd Normal Form and methods of adapting to newer technologies such as massive parallelism and NoSQL structures like Hadoop.
- Kimball (based on dimensional) - The Kimball approach is based on the use of high-performance, star schema data marts as a starting point and the data warehouse is a virtual concept consisting of the set of all data marts. This was created in response to the long timeframes many organizations were experiencing attempting to implement enterprise-scope data warehouses.

The Data Vault data modeling methodology is the most conducive to high volume, parallel loading with virtually no impact from adding new data sources.

The physical data models for ODS (REPLICA and EVOLUTION) are basically the same as FMMIS. REPLICA must be an exact match and aside from a few minor data type transformations that are needed to go from Oracle (FMMIS) to Postgres (ODS). The difference between REPLICA and EVOLUTION are as follows:

- EVOLUTION has combined paid and denied claims to help with inefficient processing of these on the FMMIS system
- The concept of PARTY from the logical data model, has been integrated into the ODS EVOLUTION model
- Claims and Encounters have been split out into different tables in ODS EVOLUTION

The Analytical Data Store (ADS) data store is a dimensional model in Redshift to allow for additional analytic capabilities beyond what the ODS structures will allow.

3.1.12.4 Elements in Rule and Policy Data Modeling

Rule and policy data modeling is the modeling of information about business rules and policy data used by rules engines, data validation services, and validation engines. The modeling strategy is largely determined by the tools used to implement these features. The elements in rule and policy data models include:

- Data edits
- Data validations
- Data translations
- Policy data
- Lookup data
- Global settings data

Policy data is organized in a Business Domain model, which groups business objects (represented by Classes) involved in a process or application. A Rule Flow defines the tasks (as Rule Task elements) associated with the process as a whole or specific objects in the process.

Lookup data is reference data that may provide multiple factors to define a profile of the transaction types applicable for a ruleset.

Global settings data are generally name / value pairs of information that change over time.

3.1.12.5 Elements in Experience Data Modeling

Experience data modeling is the modeling of information about interactions of recipients and providers in pursuing or providing a service. This is not a formal modeling process but is important to the success of FX service delivery. The specific data elements would be found in

the models outlined above. This type of modeling supports collection and analysis of behavior determinant data. The elements in experience data models include:

- Actors (refer to the many roles that may be played as part of healthcare services such as providers, recipients, case workers, judges, etc.) and their expectations in the interaction to secure/provide healthcare services
- Interaction channel, the various methods actors use such as email, social media, direct mail, mobile apps, in-person or application systems.
- Frequency, time spent by the actors, and duration of the interaction
- Impact of the interaction on the overall journey of the actor(s)

3.1.13 FX Data Repository Tool Recommendation

The *T-1: Data Management Strategy,* Section 7.4.3 *Modeling Tool Recommendation* defines the strategy for the FX data modeling tool. The FX strategy is to standardize on an enterprise data modeling tool to be used by SEAS, IS/IP, and EDW Vendors. FX Data Dictionary content (e.g., required fields, constraints, valid values, etc.) resides in a central repository and can be exported through reports, Excel, PDF, and HTML formats. Erwin has been selected as the data modeling tool for the EDW physical data models. The data dictionary will be linked within the IGC tool. IGC is the central repository for current state view.

The *T-1: Data Management Strategy,* Section 8.3 *Enterprise Metadata Management Tools* recommends the use of a centralized metadata management tool for the storage, maintenance, and dissemination of metadata. The FX projects will use a Metadata Manager to discover, collect, and centralize metadata from multiple sources across the Agency including applications, databases, data models, extract, transform, and load (ETL) tools, extract, load, and transform (ELT) tools, and BI tools. Because of the relationship between the FX Data Dictionary and metadata repository, some content will reside in both tools unless a consolidated suite is employed. IGC has been selected as the metadata management tool.

Strategic Topic 3-3: Data Dictionary Format / Tool Strategy describes the recommended format/tool the FX Data Dictionary content should be maintained in and provided to information consumers.

Data Dictionary Format / Tool	Timeline				
	Current	**2018**	**2020**	**2021**	**2024**
AHCA IT	Various (Excel, Word, etc.)	->	Agency Choice	->	->
FMMIS	iTrace/Erwin	->	Agency Choice	->	->

Data Dictionary Format / Tool	Timeline				
FX		Data Modeling Tool with export to centralized Data Management Depository	->	Erwin/Data Modeling Tool IGC Metadata Management Repository	->
Module Specific			Vendor Choice	->	->
Other Agencies	Agency Preference	->	Agency Choice	->	->

Strategic Topic 3-3: Data Dictionary Format / Tool Strategy

3.1.14 FX Data Repository Content on the Data Management Depository

The benefits of publishing content to the Data Management Depository include:

- All changes to the data models or data dictionary in the Agency data modeling tool will be published to confirm alignment between the FX Data Repository and underlying data source
- The Data Management Depository provides versioning of published FX Data Repository exports

3.1.15 FX Data Repository Changes

Requestors seeking to update or incorporate new content in the FX Data Repository, including updates to data models and data dictionary content, will follow the change control process described in the Change Management Plan (found in FX Project Management Plan) for data related changes.

The change control process for managing changes is documented in the Change Request Management process definition located in the FXPR at FX Hub > Process Definitions > Process Category: Change Management.

Strategic Topic 3-4: FX Data Repository Maintenance Process describes who implements changes to the FX Data Repository.

Data Repository Maintenance Process	Timeline				
Alternatives	**Current**	**2018**	**2020**	**2022**	**2024**

DATA REPOSITORY MAINTENANCE PROCESS	TIMELINE				
System maintenance teams update with some inconsistency in timing of Data Dictionary updates	X	X			
Identifiers request changes; approved changes implemented centrally			IS/IP / EDW /FX Modules	->	
Any identifier can make changes					
Analysis	Currently, the Agency has multiple data dictionaries (FMMIS, AHCA IT) with varying levels of completeness, discipline, frequency, and processes in terms of governance, review, and updates. The future state strategy is to create and update the FX Data Repository as the Agency goes through modular implementation. The FX Data Repository will be housed in a data modeling tool as data models and the underlying data dictionary are created and maintained. Any significant, complex, or potentially controversial changes to the FX Data Repository after modular implementation begins will go through the Data Governance Workgroup approval process. The identifier will request the change through the Project Module Control Board (MCB) who will refer those to the Data Governance Workgroup. The Data Governance Workgroup will evaluate and approve or reject the proposed change. Following approval of a change, the change will continue through the change review process as documented in the Change Request Management process definition.				

Strategic Topic 3-4: FX Data Repository Maintenance Process

3.1.16 DISTRIBUTION OF FX DATA REPOSITORY

After the SEAS Data Architect makes data model and data element updates, the SEAS Data Architect exports from the data modeling tool to the Data Management Depository. The exports provide users and consumers of FX Data Repository information access to the latest data definitions without needing licensed copies of the data modeling tool.

Strategic Topic 3-5: Distribution of FX Data Repository defines the strategy for distributing the FX Data Dictionary updates.

DISTRIBUTION OF FX DATA REPOSITORY	TIMELINE				
Alternatives	**Current**	**2018**	**2020**	**2021**	**2024**
iTrace, application documentation, database extended properties	X	X	X	X	X

DISTRIBUTION OF FX DATA REPOSITORY	TIMELINE				
FX Data Repository is only available for technical team with data model tool license				Erwin and IGC are available	->
Data Management Depository		Data Modeling Tool (2019) with export to Data Management Depository	->	Erwin will export to IGC	->
Data Management Depository, application user and technical documentation, etc.				IGC	->
Analysis	Currently, access to the Agency's multiple data dictionaries is provided through several applications (iTrace, application documentation, database extended properties, etc.). The future state strategy is to maintain the FX Data Repository in one central location. The FX Data Repository will be created in the data modeling tool with an export to the Data Management Depository. The Data Management Depository will allow anyone with access the ability to view the FX Data Repository. This will enhance collaboration and organizational empowerment so both business and IT users have consistent views of data relevant to their roles.				

Strategic Topic 3-5: Distribution of FX Data Repository

3.1.17 FREQUENCY OF UPDATES TO THE FX DATA REPOSITORY

After a data model/dictionary change is approved, the SEAS Vendor Data Architect will implement the required changes. The data modeling tool automates most of the effort. This process will make sure the FX Data Repository provides an accurate representation of the various data models and data dictionary content.

Strategic Topic 3-6: Frequency of Updates to the FX Data Repository describes how often the FX Data Dictionary receives data definition updates.

FREQUENCY OF UPDATES TO FX DATA REPOSITORY	TIMELINE				
Alternatives	**Current**	**2018**	**2020**	**2022**	**2024**
None – FX Data Dictionary is not maintained					
Data dictionary updates are ad hoc	Some AHCA IT	Some AHCA IT			

Frequency of updates to FX Data Repository	Timeline				
Periodic (e.g., monthly, bi-annual)					
FX Data Repository updates align to all production changes	FMMIS	->	FMMIS & FX Modules	->	
Analysis	Currently, Agency data dictionaries are updated on an ad hoc basis. The future state strategy is to align the updates of the FX Data Repository to all production changes. Any significant, complex, or potentially controversial changes to the FX Data Repository will go through the Data Governance Workgroup approval process. The identifier will request the change through the Project Module Control Board (MCB) who will refer it to the Data Governance Workgroup. They will evaluate, approve, or reject the proposed change. Following approval of a change, the change will continue through the change review process as documented in the Change Request Management process definition.				

Strategic Topic 3-6: Frequency of Updates to the FX Data Repository

3.1.18 Access to the FX Data Repository

The FX strategy is for the FX Data Repository to serve as the single, authoritative source for FX data definitions allowing all users and consumers of FX data to have access to appropriate content.

Strategic Topic 3-7: Access to the FX Data Repository defines the strategy to provide organizations access to the FX Data Repository.

Access to the FX Data Repository	Timeline				
	Current	**2018**	**2020**	**2022**	**2024**
Operational Data Store (ODS)	System Owners	->	->	Users of ODS (system users)	->
Information Exchange Data	Exchange Partners	->	->	Users of FX data (system users)	->
Analytic Data	Authorized data consumers	->	->	End Users by business area and persona	->
Rule & Policy Implementation Data	System Specific	->	Public	->	->
Experience Data	Unique to Business Area	->	Users of FX data	->	->

Access to the FX Data Repository	Timeline
Analysis	Currently, access to data definitions within the Agency are handled on a system by system basis, where available. As the Agency is evolving to modular implementation, the strategy is to have data definitions widely available to consumers of FX data but limited based on the specific type of data being consumed, as specified above. This limitation is for security purposes as the FX Data Repository could be used to plan a security breach in the hands of a malicious actor. The future state strategy is to have the data definitions in one central location to create a single authoritative source. This will reduce misinterpretation and ambiguity while directing consumers and users to useful data.

Strategic Topic 3-7: Access to the FX Data Repository

3.1.19 Data Elements Documentation Strategy

The FX strategy is to define a comprehensive, tool-based FX Data Repository for all FX related data contained in the ODS, Reporting Data Store (RDS), Data Warehouse, and Data Marts. The Data Management Depository will also store all data definitions and metadata used in the exchange of information through the Enterprise Service Bus (ESB).

Though there are many interfaces between FX applications, the Agency uses a small portion of the available data from other agencies. To support cross-agency data discovery, the FX strategy is to maintain a reference list and access to available data dictionaries from other agencies. As the SEAS Vendor, IS/IP Vendor, and FX Project Owners identify additional data to include in cross-agency data exchanges, the IS/IP Data Architect adds data definitions to the Data Management Depository NIEM extensions, and the SEAS Vendor Data Architect updates FX data models, if applicable.

Strategic Topic 3-8: Data Elements Documentation defines the data definitions in the FX Data Dictionary.

DATA ELEMENT DOCUMENTATION	TIMELINE				
Alternatives	**Current**	**2018**	**2020**	**2022**	**2024**
Data elements documented on a system by system basis	X	X			
Any Agency owned and controlled data that is stored in a database (e.g., ODS, RDS, Enterprise Data Warehouse, Data Marts, etc.)				X	->
Data and Metadata used in the exchange of information through the Integration Platform				X	->
Data in external sources or agencies not owned or controlled by the Agency				Consider documenting	->
Modular implementations where vendor COTS solutions bundle system functionality with proprietary data schemas and solutions					
Software as a Service solution					
Platform as a Service solution					
Analysis	Currently, the Agency documents data elements on a system by system basis. The future state strategy is to document the data elements that align to the Agency's approach to modular implementation and Integration Services. The data that should be modeled includes any Agency data owned and controlled in a database (e.g., ODS, RDS, Data Warehouse, Data Marts, etc.) and data and metadata used in the exchange of information through the Integration Platform. The Agency will evaluate documenting data from external sources or agencies to provide visibility to all data used by the Agency.				

Strategic Topic 3-8: Data Elements Documentation

3.1.20 DATA ELEMENT HISTORY

The history of data elements is an asset to help understand interpretations and policy changes over time. For longitudinal data analysis, the data element definition history helps understand

the validation rules and valid values and format of data at the time of collection. Currently, the Agency has multiple data dictionaries that use various tools and formats. AHCA IT documents data definitions in several formats (Microsoft Excel, Word, and extended database properties). FMMIS data definitions are documented in the iTrace system. Valuable data element definition history resides in these formats and tools. Throughout the life cycle of modular implementations, the creation and maintenance of the FX Data Dictionary will migrate from its current form to a centralized FX Data Repository. As part of the FX Data Repository centralization, the SEAS Data Architect will migrate data definition history from current tools to the FX Data Repository. Afterwards, data definition history will reside in the FX Data Repository. For the period of the FX modular implementations, interested stakeholders will have read only access to the current data element definitions. Upon the completion of modular implementations and certifications, the strategy is to retire the existing data dictionaries that contain historical data element definitions.

3.1.21 FX Data Dictionary Structure

The specific data dictionary structure will be predicated on the tools provided by the EDW Vendor with modifications/customizations as required by data governance. The structure shall include, at a minimum, data element names, definitions, data types, formatting rules, sensitive data identification, retention standards, and the applicable standard and attributes for each data element.

3.2 Guidelines for Data Implementation

In addition to standard data operations such as CRUD (Create-Read-Update-Delete), there are special circumstances that need to be addressed by data managers. The following sections outline guidance and standards for specialized cases.

3.2.1 Date Time Formats

Date formatting falls into two categories: storage/data exchange representation and display formatting. Date and time displays shall conform to current ISO 8601 standards to prevent ambiguous formatting (e.g., 2014-01-05 is defined in ISO 8601 as represented by YYYY-MM-DD and is interpreted as January 5 and not May 1).

3.2.1.1 Date Time Storage and Information Exchange Format

Storage and information exchange representations for each date element will format dates in one of three contexts:

- *Dating and timestamping of an event occurrence at an exact moment in time* – requires that events be recorded on a consistent timeline for proper sequencing without regard to differences in time zones or Daylight Savings adjustments. Timestamps such as these should always be recorded in Coordinated Universal Time (UTC) time. This is the default format for date storage and data exchange of transactional events.

- *Dating without regard to time* – used for dating where time is not relevant (e.g., enrollment date). This is typically recorded as a date-only data type or date/time with time set to 00:00:00.0xxx (to whatever fractional second precision is used). Using a date/time when time is truly not used can cause complications in date math and logical comparisons. If a date-only field is not an option, it is better to use 12 noon rather than midnight to represent a date because it is unaffected by time zone/Daylight Saving adjustments except in rare cases that should not be relevant to this environment.
- *Dating which requires local time knowledge* – used for dating where location-based scheduling is involved. For example, suppose an event is always scheduled for 5:00 p.m. regardless of whether the current time is standard or Daylight Saving and may or may not be relative to the specific location's time zone (i.e., should it be at 5:00 p.m. in every time zone?). Use of only UTC time would not accurately record this because UTC does not account for time adjustments. In this case UTC combined with time zone offset can be used in most cases. For calculations that require historical time adjustments, it is necessary to reference the IANA (Internet Assigned Numbers Authority) Time Zone Database that records when changes have occurred and provides the correct offsets.

Services for storing and retrieving persisted data will employ a dating format as defined above. The FX Data Dictionary will define the specific date format that is being used for any single date attribute. Data exchange services will use canonical model definitions for data interchange.

3.2.1.2 DATE TIME DISPLAY FORMAT

Display presentation format for dates should use a locality neutral format based on ISO 8601 (which uses YYYY-MM-DD). Transformations for local display formats need to clearly identify the format being used to allow for proper interpretation and ordering of data.

It will be the responsibility of the User Interface (UI) developer to ensure proper display formatting including adjustments for time zone/Daylight Savings and agency specified adjustments (e.g., display all dates for a specified use in Eastern time regardless of actual location time zone).

3.2.2 BI-TEMPORAL TRACKING

A data value, whether a single element or set of elements taken as a unit, is a recording of a fact. However, the truth and validity of a fact is either time dependent (known at a specific moment), or time constrained (known during some periods but not always). Time tracking must ensure that contradictory facts cannot be recorded (e.g., a person can't be eligible and ineligible at the same time so the change from one to the other must not overlap in time). This is not to say that multiple facts around a specific topic can't happen simultaneously (e.g., a person could be recorded as having multiple races but each one is recording a different fact without contradiction). It is critical that data element definition clarifies the issue of single vs. multiple facts simultaneously.

The FX Project Owner will implement this concept by using effective dating with additional data elements to record the time period during which the fact is known to be true for any data store containing historical data (EDW, Data Marts, etc.). All data values must have a business effective date range (begin and end) defined and temporal integrity must be maintained. Business effective dates represent the time period in which the business maintains that the fact is true regardless of when the data is entered into the system and is complicated by allowing back-dating of facts. Ranges must be used rather than just start dates to also account for non-contiguous intervals such as a recipient being eligible for some period, then not eligible and then eligible again. Care must also be taken to provide accurate results to queries for a period in which there are not known facts. For example, what was the recipient's address during the period in which they were not eligible? Unless that fact is explicitly tracked, the only valid answer is Unknown.

Care must be taken to distinguish effective dating from date-type data elements such as those in **3.2.1.1 Date Time Storage and Information Exchange Format.** Effective dating must always be done using UTC date-time to maintain a consistent timeline and to allow for accurate date calculations.

In addition to business effective dates, the FX Project Owner must also track system effective dates. These are the dates the data was either loaded into the system (begin) or has been superseded (end) by either a new data value being loaded or tracking of the data is stopped.

Tracking both sets of effective dates is known as bi-temporal tracking. Following is additional background and guidance on complex bi-temporal tracking use cases that can impact the EDW Vendor's design decisions.

Bi-temporal tracking provides for two critical use cases outside the normal business process. The first case is a requirement to be able to reproduce data from the past as it existed at that previous point in time (e.g., re-run a report that shows the data as it existed at a known point regardless of subsequent changes). The second case is the need to insert facts at a prior point in an existing timeline. This typically occurs because of back-dated changes. This allows for re-processing a report to show how the data would have looked at that time had subsequent changes already been made. Although many examples can be found, the key component is the need to either show data as it *existed* at a point in time or how it would have looked if later changes that *should have existed* had already been made.

The primary use of end dates (business and system), besides non-contiguous intervals, is to improve performance. In queries of the form "what was the data on a particular date," it is more efficient to use a BETWEEN clause in the same row than to compare dates across rows. Alternatively, with queries of the form "what is the most current data," it is more efficient to look for an indexable value far in the future (e.g., December 31, 9999) as the end date than to look for NULL or use MAX functions on the begin date.

Bi-temporal tracking using UTC date-time must be used for any data store that contains historical data. The granularity of tracking is based on data elements rather than records, which may require physical partitioning. Temporal integrity must always be maintained.

3.2.3 DATA PROTECTION

FX Project Owners must implement data controls that comply with federal, state and industry standards for protection of PHI (Protected Health Information), PII (Personally Identifiable Information), and Agency defined confidential/classified information to prevent unauthorized disclosure. This applies to any data handled by an FX Project Owner including, but not limited to, data persisted in temporary or permanent files/databases, auditing and logging contents, and data exchange message contents.

Three critical truths about data security must be addressed to ensure adequate protection:

- Data will go places that data owners don't know, can't control, and can't trust
- Encryption alone is not enough
- Detailed, comprehensive visibility into who accesses the protected data, when and how often is critical

Although enterprise security controls are addressed in the *T-8: Enterprise Data Security Plan*, FX Project Owner data protection must ensure that data is protected at rest, in use, and in transit to the extent possible. The data protection strategy must start with protecting data at the point of origin and must be independent of the device, application, or network.

Encryption alone cannot fully protect data because sharing the data with legitimate users essentially makes them co-owners of the data with the creator. Controls which prevent disclosure of that data must persist with the data and enable the originator to dictate what the recipient can do once the data is accessed (e.g., view-only without copy/paste, print or save capabilities). This persistence is critical to tethering the data to the originator because the data may end up in places the originator doesn't know about or have control over. These controls should also address platform ubiquity to include mobile and cloud and must support revocation of access should the data be located on lost or stolen devices.

Controls must include comprehensive visibility into data access, both authorized and unauthorized, and provide non-repudiation and audit information. Ideally, this information should be provided to the Agency's security information and event management (SIEM) tools for broader analysis of possible malicious intent.

Finally, administrative controls must be enacted for proper destruction of data when hardware is disposed of, or cloud vendors change. Reverse procurement and data wiping before delivery to a recycling vendor who provides auditable data destruction is recommended for IT equipment. Universal encryption of all cloud stored data with crypto shredding is the only currently reliable method of destroying cloud storage data. Crypto shredding involves encrypting the keys used to encrypt/decrypt the data in normal use and then destroying the keys used for that second encryption. This eliminates the ability to obtain the data encryption keys. Key access should always be tightly controlled external to all uses of the keys.

3.2.3.1 DATA CLASSIFICATION

Classification is a process of consistently categorizing data based on specific, pre-defined criteria so that data can be protected. Data classification schemes are typically based on one of three approaches:

1. Content-based – answers the question “What is in the data?”
2. Context-based – answers the question “How is the data being used?”
3. User-based – relies on the data owner’s knowledge and discretion to flag sensitive data.

The Agency has defined an internal data classification policy using a hybrid approach combining all three methods above that will be applied to data elements and recorded in the FX Data Dictionary. Classification is the responsibility of the FX Data Owner and is monitored for compliance by the appropriate FX Data Steward. In addition to internal classification, the Agency must ensure compliance with Health Insurance Portability and Accountability Act of 1996 (HIPAA) and HIPAA applicability will also be recorded in the FX Data Dictionary.

The Agency has identified three (3) classification levels of data:

4. Restricted – Documents are subject to compliance restrictions (e.g., PII, PHI, HIPAA) or internal classification and are not to be distributed externally unless under specific conditions. An unauthorized disclosure, compromise, or destruction would result in severe damage to the Agency, employees, vendors, or participants in MES programs.
5. Confidential – Documents are not to be distributed externally unless under specific conditions. An unauthorized disclosure, compromise, or destruction would directly or indirectly have an adverse impact on the Agency, employees, vendors, or participants in MES programs.
6. Public – Documents are acceptable for public use without restrictions. Knowledge of this information does not expose the Agency to financial loss or jeopardize the security of information assets.

The Agency must classify all data maintained by the Agency or its agents and protect the data according to classification. In the case of compound data components, the classification will be the highest-level that applies to any component part.

3.2.3.2 ENCRYPTION METHODS

The foundational technology used for key management must use industry standards such as X.509 certificates and Public Key Cryptography Standard (PKCS) and conform to best practices such as:

- Decentralized encryption/decryption processes

- Centralized key management with distributed execution
- Support for multiple encryption methods
- Centralized key user profiles for authentication and least privilege enforcement
- No decrypt/re-encrypt in case of key rotation or expiration
- Comprehensive logging and audit trails

3.2.3.2.1 DATA AT REST

All persisted data stores (databases, files, etc.) must be encrypted with a solution adhering to FIPS 140-2 certification for cryptographic modules. At a minimum, AES-128 or higher must be used for encryption.

3.2.3.2.2 DATA IN TRANSIT

All data, regardless of classification, must be transmitted securely. All transmission of web interface data shall be over SSL (Secure Sockets Layer) using TLS (Transport Layer Security) 1.2 or higher and conforming to *NIST SP 800-95: Guide to Secure Web Services* and *NIST SP 800-52 r1: Guidelines for the Selection, Configuration, and Use of TLS Implementations*.

Covered data (which includes PHI – HIPAA covered data, PII, PFI or Agency-defined classifications) that is transmitted over email must be secured using cryptographically strong email protocols such as PGP or S/MIME. Alternatively, the sender may use compliant strong file encryption and send the encrypted file as an attachment. Key exchanges must occur over a separate, secured transmission medium.

Non-web interface transmission of covered data (e.g., file uploads) can be accomplished with application-level encryption (e.g., SFTP, SCP, etc.) and must use FIPS compliant cryptographic algorithms with strength comparable to AES-128 or higher. Where non-web interface transmission cannot be accomplished with application security (e.g., local bulk updates), network-level security such as VPN tunneling, IPSec, or SSH shall be used with appropriate FIPS compliant cryptographic algorithms.

3.2.3.2.3 DATA IN USE

Data in use also needs protection but current encryption methods require decryption of the data before it can be used. Homomorphic encryption techniques are under development but not practical for widespread use. These leading-edge approaches that are not widely known but should be evaluated periodically. Organizations such as Duality, Vaultive (now part of CyberArk), and IBM have been pursuing this issue, but focus has largely been on cloud security.

Techniques like secure memory enclaves and Intel ® Software Guard Extensions (SGX) are currently the best approach to securing data in use but must be implemented in application code. IBM Cloud and Microsoft Azure have announced availability of secure enclave services

in their offerings. This will be relevant to FX Project Owners that implement cloud solutions and should be used if practical.

3.2.4 NAMING STANDARDS

Having a naming standard and applying it consistently is more important than the details of what that standard entails. Variability and inconsistency lead to confusion, errors and lost time. This becomes even more pronounced when multiple data sources are involved. The FX Naming Standard is defined in the *T3: Data Standards, Appendix D Naming Standards*.

3.2.5 CODE LOOKUP STRATEGY

The FX Code Lookup Strategy addresses: source, type, and logical/physical design efficiency. The first two deal with how code values are defined and used, and the last deals with performance and maintenance.

3.2.5.1 SOURCE

Although code values were originally intended for space savings on scarce, expensive storage and primarily used by developers, they have morphed into a common business element. Codes are now chiefly used for standardization/compliance, and only secondarily for storage savings (predominantly in data warehousing applications with large numbers of data rows). Code structures like ICD-10 (International Classification of Diseases, 10th revision) are designed to standardize a business vocabulary to ensure clear and accurate communications. USPS (United States Postal Service) state abbreviations codes and ISO 3166 country codes are examples of codes used for compliance with international and US mail processing or Internet domain name construction rules.

The FX Project Owner, in conjunction with the SEAS Vendor and Data Governance Workgroup will identify the authoritative source of values to be used for a specific code set. For all code sets that originate with an FX Project Owner, an FX Data Owner and Data Steward will be identified, and a maintenance schedule established. These code sets will be housed in the EDW, accessible through IS/IP hosted services, and referenced by the FX Data Dictionary. Bi-temporal tracking will be employed to include both future- and back-dated changes.

Authoritative sources external to FX, such as ICD-10, may be accessed through services hosted in IS/IP that connect directly to the authoritative source, provided that relevant Service Level Agreements (SLAs) are maintained. If performance issues cannot be resolved, the code set will be imported to the EDW and managed locally. An FX Data Owner and Steward will be identified, and a maintenance schedule established. Bi-temporal tracking will be employed to include both future- and back-dated changes if approved by the authoritative source.

3.2.5.2 DEFINITION

Code set definitions fall into one of two types: designation or classification. Designations apply a solitary label to an entity; the codes are unrelated to one another. Classifications are

hierarchical codes that imply inheritance of higher-level values and are used to aggregate to various levels. Race and gender codes are examples of designations. A vehicle taxonomy such as **Exhibit 3-4: Classification Code Structure** would be a classification type code.

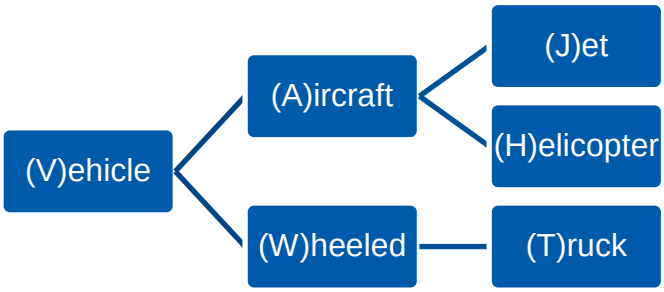

Exhibit 3-4: Classification Code Structure

All coding systems must be identified as to type to determine the minimum necessary data elements to use. Designations require only identification and description attributes; classifications require additional attributes to manage the interrelationships between members. All coding systems will employ bi-temporal tracking.

3.2.5.3 DESIGN EFFICIENCY

Ultimately all use of codes represents a substitution of exactly one value for another. One expected FX strategy for code design is to have a single structure to house all codes (e.g., a *universal* code table). The advantage is that all codes are managed in a single location and maintenance of effective dating can be handled with a single module. The disadvantages are that extra metadata is required to segment the applicable codes for specific uses, and duplicate codes can be problematic, requiring complex or surrogate keys to identify the correct code (e.g., F could be used for Female in the Gender codes, but F could also mean Fee in the Financial codes). Large code tables can exhibit performance issues, but these can usually be resolved in the physical design through use of techniques like partitioning or in-memory tables.

All code access will be through IS/IP hosted services. The underlying physical structure used will be determined by the EDW Vendor and subject to SLA compliance. The ODS is designed to be the single authoritative source of data for all FX Modules (including the Data Warehouse/Data Marts) and therefore will be the authoritative source of code values in any downstream system. Any external reproduction of code values for module processing will require an approved synchronization plan.

3.2.5.4 DATA MODELING OF CODES

There are different perspectives on whether code structures should be included in logical data models. The common reason code tables would be excluded from logical models is because code values are primarily implemented for performance reasons and belong only to the physical design. The FX strategy is to take a pragmatic approach and include those code sets

in the logical model that enhance understanding, require special maintenance, or serve a core business reason. The logical model will show discrete code structures for each individual or shared use. The physical implementation will be left to the appropriate FX Project Owner.

Physical design is the responsibility of the IS/IP and EDW Vendors; however, for code structures that are defined only in the physical layer, it is important that code history be maintained. Specifically, in the ODS and Master Person/Organization Indexes it is imperative that those codes are historically tracked. The data warehouse will maintain history of all data, including code values.

3.2.6 NORMALIZATION

Normalization is a design process to address various anomalies that can occur in relational databases because of insertions, updates and deletions and to reduce restructuring as new types of data are introduced. The FX Logical Data Model will conform to the industry standard of third normal form (3NF). The specialized forms of Boyce Codd normal form (BCNF) and elementary key normal form (EKNF) will be used if needed to deal with specific, rare anomalies that 3NF does not address.

Denormalization is used during physical design to improve query performance with a trade-off of complexity of managing data integrity, additional storage, and decrease in insert/update performance. Dimensional, data vault, and NoSQL (key-value, columnar, graph, document, etc.) physical models are significantly denormalized. The EDW Vendor will determine the need for physical implementation appropriate denormalization from the logical model.

3.2.7 STANDARDS FOR ADDRESS

Address Information System (AIS) products available from the United States Postal Service (USPS) and third-party vendors (e.g., Lob, Google, Melissa Data, Informatica, etc.) provide a variety of address services that include address validation and verification, address enrichment, authoritative data sets, component lookups, and delivery statistics. Addressing, due to the variety and complexity both within the U.S. and internationally, has been broken down into three segments: elements, templates and rendition. Elements represent the individual data components that comprise an address. Templates define how the elements are assembled for proper construction to accommodate mail delivery systems. Rendition are rules for spacing, punctuation and format.

The Universal Postal Union (UPU) (established in 1874) is the second oldest international organization in the world and the United States is a participating member. The UPU has defined an international set of data elements that have been agreed to by all member nations. Each member nation then defines local templates for formatting and rendering addresses for mail delivery. The USPS standard is based on the components of the UPU standard and primarily concerns the template and rendition. (sources: Universal Postal Union Standard S42: International Postal Address Components and Templates and US Postal Service Publication 28: Postal Addressing Standards)

3.2.7.1 ELEMENTS

S42 defines elements as the smallest meaningful parts of a name or address. The choice to use parsed data elements was to allow for the variety of addressing methods and to provide mechanisms for maintaining address quality, validity, and renditioning. It also addressed the problem of storing multinational addresses in a single database.

The UPU elements are organized into a two-tier structure that consists of four first level identifiers and 34 second level identifiers. The second level identifiers are individual data elements. The identifiers are listed in the following two subsections.

3.2.7.1.1 FIRST LEVEL

10 Addressee specification segment
20 Mailee specification segment
30 Recipient dispatching information segment
40 Delivery point specification segment

3.2.7.1.2 SECOND LEVEL

00 organization name	17 district/sector
01 legal status	19 delivery service identifier
02 organizational unit	20 alternate delivery service identifier
03 function	21 thoroughfare
04 addressee role descriptor	24 street number or plot
05 form of address	26 building/construction
06 given name	28 extension designation
07 surname prefix	29 wing
08 surname	30 stairwell
09 name qualifier	31 floor
10 qualification	32 door
11 mailee role descriptor	33 supplementary dispatch information
12 defining authority	34 supplementary delivery point data
13 postcode	35 delivery service qualifier
14 country name	41 country code
15 region	43 multi-country region
16 town	44 international routing information

In addition to these elements, there are defined modifiers for specific elements (e.g., 40.21 (Thoroughfare) also permits pre and post directional qualifiers such as E, N, S, W, NW, SE, etc. and a type designation such as ST, BLVD, AVE, etc. Element 40.13 (Postcode) in the U.S. has two parts to accommodate the original ZIP and the ZIP+4 components.

3.2.7.2 U.S. TEMPLATE

The U.S. defined template allows for several alternative layouts as defined by the UPU and USPS Publication 28. Publication 28 focuses largely on layout and rendition instructions for mailings and includes components for people and organizations that will be outside the scope of this standard.

The delivery address component defined in both UPU and USPS consists of two required lines of data and one optional. Publication 28 refers to these as Delivery Address Line and Last Line. The Delivery Address Line can be broken into two lines (Primary and Secondary) for special circumstances such as dual addresses (both street and P.O. Box), excessive length on the mail piece, and Urbanization designations for Puerto Rican addresses. The templates for the Delivery Address Line and Last Line are included in the following two subsections.

3.2.7.2.1 DELIVERY ADDRESS LINE

40.21, 40.19, 40.35	for Rural Route or Military addresses OR
40.19, 40.20	for PO Box or General Delivery addresses OR
40.24, 40.21 + modifiers, 40.32 + modifiers	for Street addresses[1] [2]

3.2.7.2.2 LAST LINE

40.16, 40.15, 40.13 (Part One and optional Part Two)

3.2.7.3 ADDRESS DATA STORAGE STANDARD

Storing the granular components of the Delivery Address Line(s) is unnecessarily detailed for the Agency's purposes. Data for this line will be stored in a single field. The contents of this field must conform to USPS Publication 28 for data components and format and should be validated as a legitimate address via AIS services. Last Line data elements will be stored as separate data components due to their use in contexts other than addressing. These data

[1] Puerto Rico addresses have an optional 40.17 used for Urbanization. Publication 28 allows it to be used on a Secondary Delivery Address Line or embedded in the Delivery Address Line.

[2] Commercial Mail Receiving Agencies (CMRA) and companies using internal mail stops can add 40.19 (PMB – Private Mail Box or MSC – Mail Stop Code) and 40.35 to the end of the Delivery Address Line or as a Secondary Delivery Address Line above.

elements must also conform to USPS standards for content and validated through AIS services, when used as part of an address.

Two additional data fields will be provided: Urbanization field to support addresses in Puerto Rico and a Secondary Delivery Address Line to accommodate those exceptions to a single Delivery Address Line provided for in USPS Publication 28.

3.2.7.3.1 SAMPLE ADDRESSES:

Puerto Rico – Separate Urbanization

Sample Address	Component Used	Component Value
URB San Joaquin	(40.17) district/sector	San Joaquin
38 Calle A Rodriguez	(40.24) street number or plot	38
Adjuntas PR 00601-2302	(40.21) thoroughfare	Calle A Rodriguez
	(40.16) town	Adjuntas
	(40.15) region	PR
	(40.13 1st part) postcode	00601
	(40.13 2nd part) postcode	2302

Puerto Rico – Embedded Urbanization

1234 URB Los Olmos	(40.24) street number or plot	1234
Ponce PR 00731-1235	(40.17) district/sector	Los Olmos
	(40.16) town	Ponce
	(40.15) region	PR
	(40.13 1st part) postcode	00731
	(40.13 2nd part) postcode	1235

Rural Route

RR 3 Box 624A	(40.21) thoroughfare	RR 3
Provo UT 84604	(40.19) delivery service identifier	Box
	(40.35) delivery service qualifier	624A
	(40.16) town	Provo
	(40.15) region	UT
	(40.13 1st part) postcode	84604

P.O. Box

PO Box 25	(40.19) delivery service identifier	PO Box
Quincy FL 32351-0001	(40.20) alternate delivery service identifier	25
	(40.16) town	Quincy
	(40.15) region	FL
	(40.13 1st part) postcode	32351
	(40.13 2nd part) postcode	0001

Street with Qualifiers

2100 N Main ST SW	(40.24) street number or plot	2100
Provo UT 84604-1253	(40.21) thoroughfare	Main
	(40.21) thoroughfare pre-directional	N
	(40.21) thoroughfare type	ST
	(40.21) thoroughfare post-directional	SW
	(40.16) town	Provo
	(40.15) region	UT
	(40.13 1st part) postcode	84604
	(40.13 2nd part) postcode	1253

Street with Door

775 Broadway APT 2F	(40.24) street number or plot	775
New York NY 10001-6315	(40.21) thoroughfare	Broadway
	(40.32) door type	APT
	(40.32) door indicator	2F
	(40.16) town	New York
	(40.15) region	NY
	(40.13 1st part) postcode	10001
	(40.13 2nd part) postcode	6315

Complex Street (Qualifiers and Door)

2800 N Los Felices CIR E UNIT C10	(40.24) street number or plot	2800
Palm Springs CA 92262	(40.21) thoroughfare	Los Felices
	(40.21) thoroughfare pre-directional	N
	(40.21) thoroughfare type	CIR
	(40.21) thoroughfare post-directional	E
	(40.32) door type	UNIT
	(40.32) door indicator	C10
	(40.16) town	Palm Springs
	(40.15) region	CA
	(40.13 1st part) postcode	92262

Military APO (Army Post Office)

USS NIMITZ CVN 68	(40.21) thoroughfare	USS NIMITZ CVN 68
APO AE 09521-2820	(40.19) delivery service identifier	APO
	(40.35) delivery service qualifier	AE
	(40.13 1st part) postcode	09521
	(40.13 1st part) postcode	2820

General Delivery

General Delivery	(40.19) delivery service identifier	General Delivery
Honolulu HI 96818-4326	(40.16) town	Honolulu
	(40.15) region	HI
	(40.13 1st part) postcode	96818
	(40.13 2nd part) postcode	4326

Complex Street (Door plus Commercial Mail Receiving Area (CMRA) exception)

10 Main ST STE 11 PMB 24	(40.24) street number or plot	10
Herndon VA 22071	(40.21) thoroughfare	Main
	(40.21) thoroughfare type	ST
	(40.32) door type	STE
	(40.32) door indicator	11
	(40.19) delivery service identifier	PMB
	(40.35) delivery service qualifier	24
	(40.16) town	Herndon
	(40.15) region	VA
	(40.13 1st part) postcode	22071

SECTION 4 DATA STANDARDS COORDINATION

4.1 DATA STANDARDS STRATEGY

Formal standards play an important role in implementing data governance, supporting decisions about data, and data management solutions. These standards guide the implementation of FX projects, key data domains, and data management capabilities. **Strategic Topic 4-1: Organizational Involvement in Data Standards** defines the organizations involved in defining and using data standards.

ORGANIZATION INVOLVEMENT	TIMELINE				
	Current	**2018**	**2020**	**2022**	**2024**
MMIS	Unique to system	->	X	->	
AHCA IT	Unique to system	->	X	->	
Cross-Medicaid Agencies		Inform	Consult	->	
Statewide		Inform	Consult	->	
Multi-state		Inform	->		
Providers and health plans	Mandate	->			
DMS Division of State Technology	Potential Input	Inform	Consult	->	
Analysis	Currently, the definition and use of FX data standards is unique to each division or business area within the Agency. The future strategy is for Agency systems owners and cross-agency Medicaid systems owners to work together in defining and using data standards. States, Providers, Health Plans, and the DMS Division of State Technology will be informed of the definition and use of FX data standards.				

Strategic Topic 4-1: Organizational Involvement in Data Standards

Exhibit 4-1: Data Standards Creation/Expansion shows the expected FX projects where data standards will be introduced or expanded.

Data Standard Types	Strategy Document Creation	Integration Services / Integration Platform	Enterprise Data Warehouse	Module Implementation (e.g., Provider, Recipient, Core Processing)
Standards authorities and governing bodies	✓			
Information Exchange	✓	✓		
Database Object Naming		✓	✓	
Code Object Naming		✓		✓
Data Modeling & Dictionary	✓(conceptual/logical)		✓(physical)	
Web Services		✓		
Security	✓	✓	✓	✓
Metadata Exchange	✓			
Privacy	✓		✓	
Messaging		✓		
Acronyms	✓	✓	✓	✓
Abbreviations	✓	✓	✓	✓
Agency Glossary	✓	✓	✓	✓
Healthcare Data Exchange (e.g., X12)		✓		✓

Exhibit 4-1: Data Standards Creation/Expansion

The *T-6: Technology Standards,* Section 4 *Technology Standards Reference Guide (TSRG)* defines technology standards and the purpose of the TSRG. The TSRG is the repository of data, project management, security, and technology standards applicable to the administration and operation of the enterprise and future state enterprise. Content for the TSRG is in an Excel list in the FXPR located at FX Hub > Reference Materials > Category: Technology, which adheres to the MITA Framework.

Standards addressing the same topic are created by different organizations and are often aligned and consistent. Higher-level organizations (typically those closest to the creation and

management of data) may adopt stricter or more specific standards. In some cases, standards may conflict, or an organization may provide guidance that certain standards are waived or not applicable. The TSRG seeks to help stakeholders understand not only the universe of applicable standards, but also to provide the structure to harmonize conflicting standards or guidance.

Exhibit 4-2: TSRG Standards Hierarchy shows the types of organizations that are sources of relevant technology standards.

Exhibit 4-2: TSRG Standards Hierarchy

Projects are to follow applicable standards about a topic by viewing standards based on the hierarchy and considering the guidance from the highest level of the hierarchy that is applicable to the project. If Agency guidance existed that overrode state, or lower levels in the hierarchy, projects would follow the Agency guidance. For a project for which CMS/Office of the National Coordinator for Health Information Technology (ONC) guidance was not relevant, standards from that level of the hierarchy would not be applicable.

4.2 DATA STANDARDS TAXONOMY

A taxonomy is a hierarchical structure separating concepts into specific classes or categories based on common characteristics. The taxonomy provides a conceptual framework for discussion, analysis, or information retrieval. The *T-6: Technology Standards,* Section 4 *Technology Standards Reference Guide* defines the guide and the taxonomy for technology, security, and data standards. Key elements of the taxonomy are the domain, area, and category. Data standards are found in the following taxonomy concepts in the TSRG in the FXPR at FX Hub > Reference Materials > Category: Technology > Technology Standards Reference Guide (TSRG):

- Data standards definitions used in the exchange of information
  - These are data standards used in the exchange of information and specify standards for both format and content.
  - Domain: Technical
  - Area: Information
  - Category: Data Standards
- Data standards setting authorities and governing bodies
  - These are the authoritative governing bodies that preside over the establishment, maintenance and deprecation of data standards.
  - Domain: Technical
  - Area: Standards Authority
  - Category: Data Standards

4.3 DATA STANDARDS COORDINATION EVENTS

The *T-6: Technology Standards, Attachment D,* Section 3 *FX Standards Communication* defines the communication strategy of the data standards related coordination events and the process to handle the event and communicate to the affected audience. The data standards events requiring communication outlined in the document include:

- Identification or development of a new FX data standard
- Modification to an existing FX data standard
- Waivers or retirement of FX data standards
- Involvement of new organizations in the FX
- Processing of feedback provided about FX data standards

Data standards are ever evolving, including the introduction of new federal, state, or agency specific standards. **Strategic Topic 4-2: New Data Standards** defines the anticipated timing of the adoption of new data standards for selected FX subject areas.

NEW DATA STANDARDS	TIMELINE				
	Current	**2018**	**2020**	**2022**	**2024**
Invoicing/financial processing		FL Treasury payment processes (2019)	FL PALM	->	
Claims processing	X12	->	Standardized response messages		
New data types				Sensor	Genetic
Social determinants of health			Data aggregation	->	
Clinical Data				Evaluate FHIR	
Experience Data			Recipient Data	Behavioral	
EHR	HL7	->			
Analysis	Currently, new data standards are evaluated and adopted on a system by system basis. The future state strategy is to expand the use of data standards and be more holistic in the Agency's evaluation and adoption of data standards.				

Strategic Topic 4-2: New Data Standards

4.4 DATA STANDARDS COORDINATION PROCESS

Exhibit 4-3: TSRG New/Updated Standards Process includes a flow chart of the steps to communicate and coordinate with affected stakeholders when there is a new or updated data standard. This process aligns with the process outlined in the *T-6: Technology Standards, Attachment D Technology Standards Communication, Support, Compliance, and Compliance Reporting Procedures*.

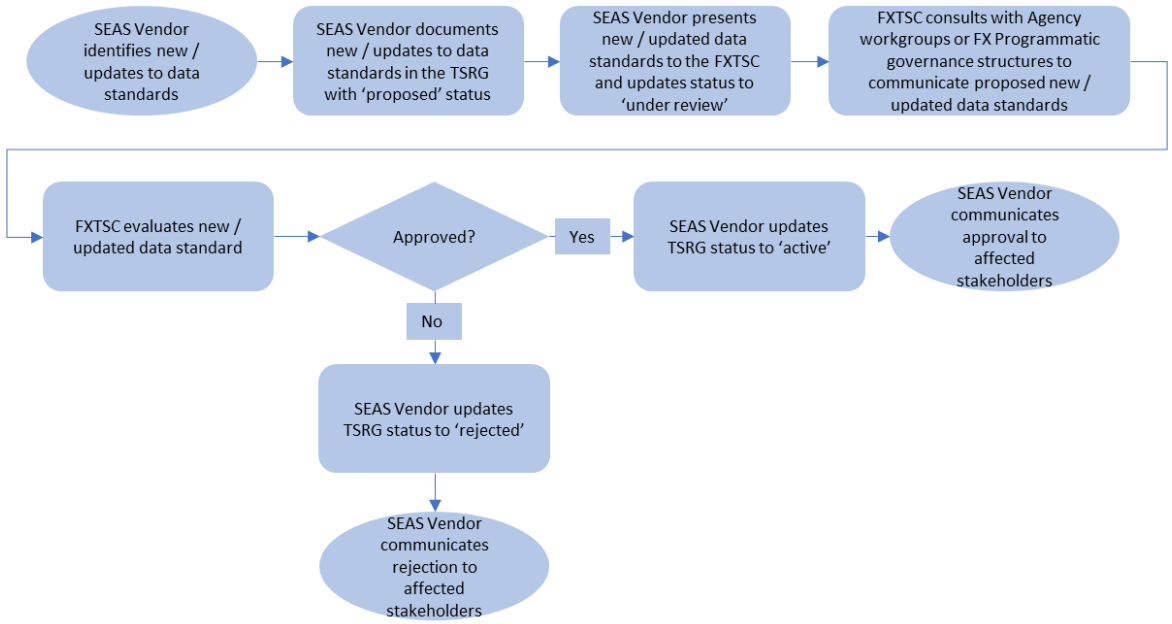

Exhibit 4-3: TSRG New/Updated Standards Process

SECTION 5 DATA STANDARDS COMPLIANCE ASSESSMENT

Periodically assessing standards compliance improves data quality and consistency. There are three main events in the FX Projects Life Cycle when the SEAS Vendor assesses compliance:

- Procurement Phase
- Design Phase
- Implementation Phase

During each event, the SEAS Vendor uses the compliance assessment approach and documentation in the *T-6:* Technology *Standards, Attachment D,* Section 5 *FX Standards Compliance Assessment* and Section 6 *FX Standards Compliance Reporting*.

ENABLING CAPABILITY	PROCUREMENTS	DESIGN PHASE	IMPLEMENTATION
Automated Compliance Tests		✓	✓
Deliverable and Work Product Review	✓	✓	✓
Key Personnel Interviews and Review Meetings	✓	✓	✓
Third Party Compliance Testing			✓
System Artifact Type Review		✓	✓
Random Sample Review		✓	✓
Vendor Compliance Attestation	✓	✓	✓

Exhibit 5-1: Standards Assessment Approach**Exhibit 5-1: Standards Assessment Approach** shows representative standards assessment(s) the SEAS Vendor performs at events in the FX Projects Life Cycle.

ENABLING CAPABILITY	PROCUREMENTS	DESIGN PHASE	IMPLEMENTATION
Automated Compliance Tests		✓	✓
Deliverable and Work Product Review	✓	✓	✓
Key Personnel Interviews and Review Meetings	✓	✓	✓
Third Party Compliance Testing			✓
System Artifact Type Review		✓	✓
Random Sample Review		✓	✓
Vendor Compliance Attestation	✓	✓	✓

Exhibit 5-1: Standards Assessment Approach

The SEAS Vendor will maintain a FX Standards Compliance Assessment Form (i.e., FX Hub > Standards & Plans > Category: Technology > Technology Standards (T-6) > Attachment E) to capture and store the results of the standards compliance assessments performed for each FX project. The compliance assessment form contains a pass/fail for compliance, the detailed results from the compliance assessment, and the corrective action necessary to achieve compliance.

5.1 PROCUREMENT PHASE DATA STANDARDS ASSESSMENT

Data standards definitions reside in the TSRG in the FXPR at FX Hub > Reference Materials > Category: Technology and are available to FX Project Owners as part of the procurement process. Procurement responses should acknowledge, accept, and demonstrate understanding of FX Data Standards compliance expectations. The Agency will evaluate if each potential FX Project Owner meets the procurement requirements which include agreement to comply with FX Data Standards and consider any FX Project Owner recommendations to update or adopt new FX Data Standards. The benefits of assessing FX Data Standards during the procurement process include:

- Governance around the vendor selection process
- Alignment on what vendors are recommending to the Agency vs. MITA and state data standards

The SEAS Vendor will refer to any new or updated data standards that result from procurement processes to the Data Standards Update Process.

5.2 Design Phase Data Standards Assessment

After the Agency selects a FX Project Owner to implement an FX project, the FX Project Owner designs the solution. The SEAS Vendor will assess the solution design to validate compliance with FX Data Standards. The benefits of assessing the FX Data Standards compliance during the design phase include:

- Confirmation that the proposed solution adheres to data standards
- Alignment on new or updated data standards that are reflected in the TSRG
- Remediating compliance issues is less expensive the earlier the issues are addressed

5.3 Implementation Phase Data Standards Assessment

After the approval of the design of the solution, FX Project Owners will implement the solution. The SEAS Vendor will perform periodic data standards compliance assessments during the implementation phase. The benefits for assessing the data standards periodically during the implementation phase include:

- Alignment that the FX Project Owners did not deviate from the data standards outlined in the design stage
- Validation that the solution meets the Agency's requirements

SECTION 6 DATA STANDARDS UPDATE PROCESS

It is the responsibility of the SEAS Vendor and the Agency to periodically update and maintain the FX Data Standards. The SEAS Vendor will perform an annual assessment of all data standards in the TSRG to determine:

- Updated versions of existing standards
- New data standards that should be adopted
- Agency needs that should be escalated to FX Technology Standards Committee

The benefits for using a defined process for updating data standards include:

- Reduced latency of adoption
- Improved data quality
- Balanced input from stakeholders
- Improved productivity and output

The *T-6: Technology Standards,* Section 4 *Technology Standards Reference Guide* describes the structure, maintenance, and communication of the TSRG. The *T-6: Technology Standards, Attachment B How to Maintain the Technology Standards Reference Guide* describes the procedures to maintain content in the TSRG. The document includes definitions of the fields in the TSRG (e.g., standards name, version, maturity, owning organization, compliance approach, status, etc.), steps for creating a new standard, and steps for updating an existing standard. The TSRG has a Compliance Approach section that contains a narrative that will be used to define the process and list of events of verifying adherence to the applicable standard.

Exhibit 6-1: Data Standards Refresh Events describes the events when the data standards will be reviewed and updated as necessary.

EVENT	DESCRIPTION
Quarterly Review	The SEAS Vendor will conduct a quarterly review of the data standards in the TSRG looking for updates to existing data standards and new data standards relevant to the Agency that should be added to the TSRG.
Issuance of ITN / Procurement	As part of the creation of ITN/Procurement documentation, The SEAS Vendor will conduct a review of the data standards in the TSRG looking for updates to existing data standards and new data standards relevant to the Agency that should be added to the TSRG.
Publication of new MITA Standard(s)	If there is a material change in MITA Part II – Information Architecture, the SEAS Vendor will conduct a review of the data standards in the TSRG as compared to MITA. If required, existing data standards will be updated and new data standards relevant to the Agency will be added to the TSRG.

Exhibit 6-1: Data Standards Refresh Events

SECTION 7 APPENDIX A – ATTACHMENTS

Attachment A – Data Naming Standards

Attachment A – *Data Naming Standards* is stored in the FXPR to serve as supporting documentation for the *T-3: Data Standards*. (i.e., FX Hub > Standards & Plans > Category: Technology > Data Standards (T-3))

Contains the standards and guidelines for naming data elements of various types.

SECTION 8 APPENDIX B - REFERENCES TO OTHER DELIVERABLES

The following attachments are stored in the FXPR to serve as supporting documentation for the *T-3: Data Standards* (i.e., FX Hub > Standards & Plans > Category: Technology)

Data Management Strategy (T-1), Section 4.1 Data Management and Data Stewardship

Describes the governance of data and the roles and responsibilities involved.

Data Management Strategy (T-1), Section 7.4.3 Modeling Tool Recommendation

Discusses the recommendation and rationale for an Agency data modeling tool.

Data Management Strategy (T-1), Section 8.3 Enterprise Data Management Tools

Discusses the features of metadata management tools and the strategy for selection and use.

Technology Standards (T-6)

Establishes the MITA compliant Florida Medicaid Technology Standards Reference Guide (TSRG) and Technology Standards Reference Model (TSRM) and describes a maintenance process.

Technology Standards (T-6), Section 4 Technology Standards Reference Guide (TSRG)

Defines technology standards and the purpose of the TSRG.

Technology Standards (T-6), Attachment B How to Maintain the TSRG List

Describes the procedures to maintain content in the Technology Standards Reference Guide.

Technology Standards (T-6), Attachment D, Section 3 FX Standards Communication

Defines the communication strategy of the data standards related coordination events.

Technology Standards (T-6), Attachment D, Section 5 FX Standards Compliance Assessment

Describes the process for performing a standards compliance assessment.

Technology Standards (T-6), Attachment D, Section 6 FX Standards Compliance Reporting

Describes the requirements for reporting the results of a standards compliance assessment to the Agency.

Technology Standards (T-6), Attachment D, Technology Standards Communication, Support, Compliance, and Compliance Reporting Procedures

Describes the processes to communicate new and modified standards, and compliance expectations, to stakeholders; support stakeholders' adherence to standards; assess stakeholders' compliance to standards; and communicate levels of standards compliance to the Agency.

Enterprise Data Security Plan (T-8)

Outlines the issues, processes and controls necessary to create and manage an enterprise-level, data-centric security plan.

PD-9 Data Naming Abbreviations Appendix N

Reflects the new naming conventions.

FXPR > EDW > Project Artifact Directory > 8-Deliverables > PD-09 System Design Document CONFIDENTIAL > 4-Final > FX-EDWI-PD-9-Appendix-N-Physical-Data-Model-Naming-Standards-100.docx

HQA Data Mart Acceptable Abbreviations

Outlines additional naming standards/abbreviations.

FXPR > EDG > Project Artifact Directory > 3-Execution > Naming Standards and Data Conversion Clean Up > HQA_DataMart_AcceptableAbbreviations_ForFXConsideration.xlsx