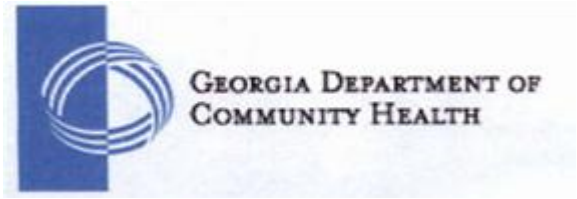

Georgia DCH Medicaid Enterprise – Data Management Strategy

February 2015

Version	Date	Document Revision History	Document Author/Reviser
1.0	11/21/2014	Draft Deliverable Submission	PCG
2.0	1/22/2015	Updates based upon collective feedback from DCH review	PCG
3.0	2/27/2015	Finalized and approved by DCH	PCG

Introduction

This Enterprise Data Management Strategy (EDMS) Document has been developed for Georgia Medicaid to further support their efforts to implement Information Architecture across the enterprise. The document sections are designed to provide the following information:

Data Strategies and Usage

DCH Strategic Plan

An effective Enterprise Data Management Strategy will provide critical support for many of the key goals and strategies identified in the Georgia DCH Strategic Plan. The Strategic Plan clarifies the importance of the role that DCH plays in promoting health and providing quality health care to Georgians.

Mission Statement

The mission of the Department of Community Health is to provide access to affordable, quality healthcare to Georgians through effective planning, purchasing, and oversight.

Vision Statement

Georgia Department of Community Health is committed to a lean and responsive state government that promotes the health and prosperity of its citizens through innovative and effective delivery of quality health care programs.

Strategies and Initiatives

There are several strategies and initiatives identified in the Strategic Plan that should be considered during the development and adoption of the EDMS. A few examples include:

- Examine Lean Six Sigma as an approach to improve performance, service delivery, accountability, and to achieve greater efficiency.
- Implement basic foundation to connect Qualified Entities enabling the secure exchange of patient health data for the purposes of treatment.
- The Integrated Eligibility System (IES) will bring together the various benefit programs provided by the State (Medicaid, TANF, Food Stamps, WIC and Child Care).
- Update Information Privacy Security Incident Response procedure and IT policies.

An effective Enterprise Data Management Strategy will support the consistent and efficient management of these and many other initiatives. Many initiatives will require easily accessible and trusted data to measure results. The adoption of the EDMS enables this capability.

Health Information Exchange

Health Information Exchange (HIE) concerns the mobilization of healthcare information electronically across organizations within a state, region, community or hospital system. The goal of HIE is to facilitate access to, and retrieval of, clinical and administrative data to provide secure and more timely, efficient,

effective, and equitable patient-centered care. HIE also directly supports public health reporting to county, state and federal authorities which assists in the treatment, monitoring and control of disease in addition to analysis of the health of their populations. To achieve these goals, HIE initiatives target improving technology, interoperability, standards utilization, and harmonization of business information systems.

The DCH Strategic Plan identified several key initiatives aimed at advancing HIE in Georgia. Some examples are:

- Establish connection between existing local and regional HIE networks and other large health systems within Georgia.
- Facilitate the acquisition of certified Electronic Health Record (EHR) technology by providing resources to connect to the state Health Information Exchange.
- Deploy secure technology solutions to establish interoperability between the Georgia HIE and specific statewide functions, connectivity with specific states in the Southeast, and the federal bridge and national Health Information network (NwHIN).
- Implement basic foundation to connect Qualified Entities enabling the secure exchange of patient health data for the purposes of treatment.

These are all current opportunities that DCH has to support data integration, health information exchange and use of clinical data in addition to administrative data to support advancement of Medicaid Information Technology Architecture (MITA) maturity. As DCH progresses in obtaining common data standards and practices within the Medicaid system, it will need to coordinate data management strategies with other HIE entities, both within state as well as in the wider Georgia healthcare community.

In August of 2013 Georgia implemented the Georgia Health Information Network (GaHIN). The GaHIN is considered a 'network of networks' rather than a single database which brings together disparate sources of electronic healthcare data. Currently Georgia Medicaid is a supplier of data to the network and is not yet retrieving data from it. However, a key finding of the 2014 SS-A was the need for the types of data that can be obtained from the GaHIN. As the EDMS is implemented coordination with the GaHIN will be necessary to support the business and data goals of the Enterprise.

Data Analytics and Business Intelligence

The Department's Decision Support Services (DSS) Team, housed within the Division of Information Technology is responsible for data reporting and analytics. The team uses multiple systems, which include the MMIS reporting repository administered by HP and a Decision Support System and Data Warehouse (DSS/DW) supported by Truven Health Analytics. All reporting systems include access to Medicaid enrollment, medical and pharmacy data for both Fee-For-Service and Managed Care populations. With an objective to provide data to assist with improved business performance and to make better informed decisions, the DSS has started an initiative to procure an enterprise data warehouse solution (EDS) with business intelligence (BI) and predictive analytic capabilities. Implementation of the data warehouse with alignment to the EDMS is expected to enable business processes which create an intelligence strategy driven by business objectives and prospective outlooks which will support better decision-making capabilities and assist DCH in achieving its desired goals.

Background

The Medicaid Information Technology Architecture (MITA) Framework Information Architecture (IA) recommends data standards and identifies enabling technologies and interoperable designs for data exchange. To achieve this goal, the Georgia Department of Community Health (DCH) MITA initiative requires a strategy for identifying, designing and managing enterprise data across its information systems. **Managing data across the enterprise requires strategic planning, coordination and adherence to a shared set of principles** to achieve the goals of interoperable systems and greater MITA maturity. Methods and activities specified in this deliverable that are involved in achieving the strategy are based on the federal Centers for Medicare & Medicaid Services (CMS) MITA 3.0 framework for IA.[1]

The MITA IA has a mandatory relationship with both the MITA Business Architecture (BA) and the MITA Technical Architecture (TA) frameworks. The BA describes the business processes along with data input, data output, and shared data required. The TA describes the technology enablers associated with different levels of maturity. The IA provides the bridge between the business view of the information and the technical view of the data.

Context of Enterprise Data Management within the MITA Framework

The MITA IA components leading to information capability are:

- Data Management Strategy (DMS) – Provides a structure that facilitates the development of information/data, effectively shared across a Medicaid Enterprise to improve mission performance.
- Conceptual Data Model (CDM) – Represents the overall conceptual structure of the data, providing a visual representation of the data groups needed to run an enterprise or business activity.
- Logical Data Model (LDM) – Identifies all of the logical data elements that are in motion in the system or shared within the Medicaid Enterprise.
- Data Standards – Discusses the available data standards, the benefits of data standards, and using them.
- Information Capability Matrix – Defines the information capabilities used in a business process and informs the identification of technical capabilities

MITA describes these components as living models that will evolve throughout the MITA life cycle. The MITA Framework tailors the level of detail in each model to meet the specific needs of the intended audience. Figure 1 provides an overview of the components of the IA.

[1] CMS MITA Framework version 3.0, Part II – Information Architecture, Chapter 6 – Information Capability Matrix, page 5, dated February 2012.

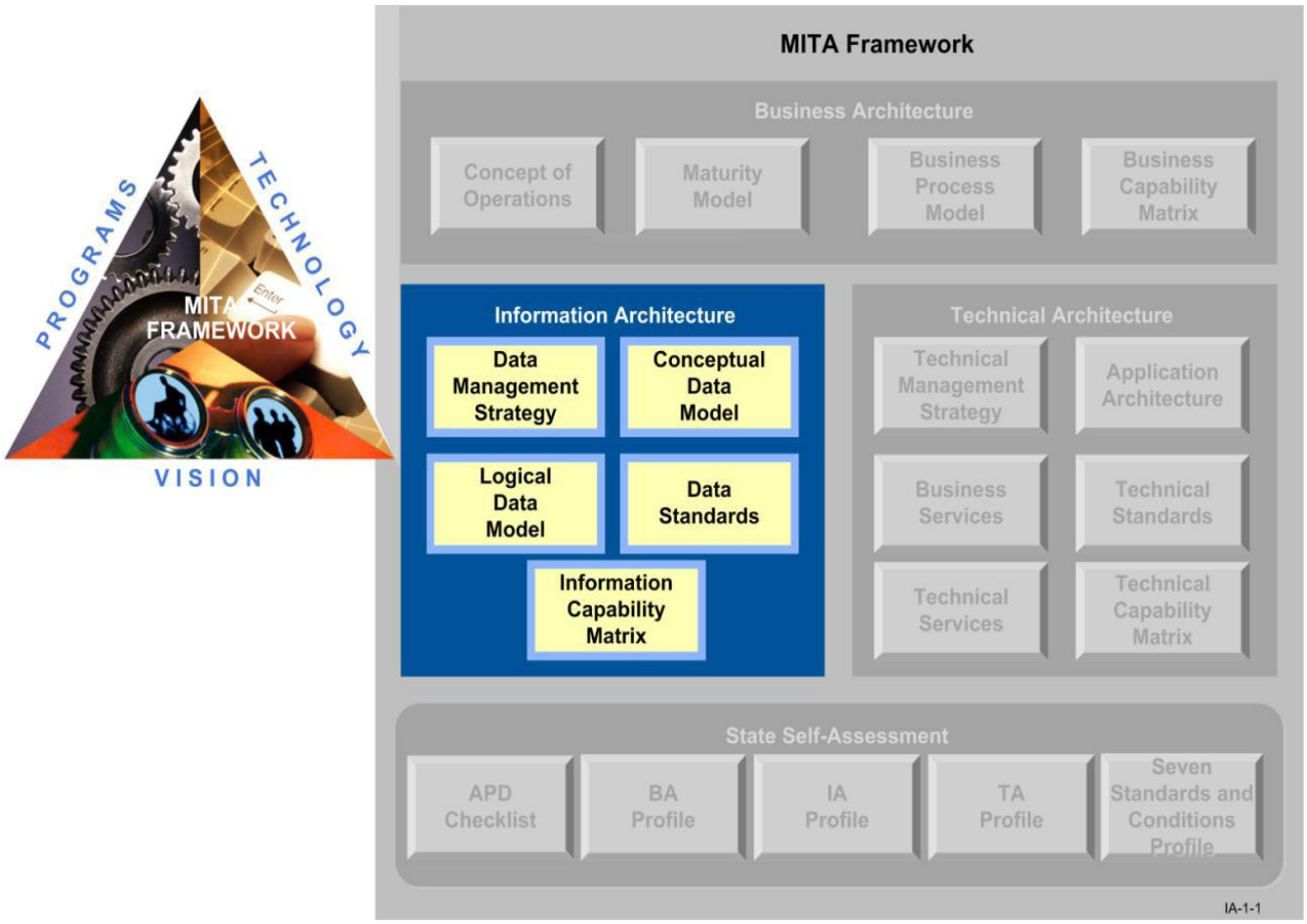

Figure 1: MITA Information Architecture Components

Purpose

The EDMS is a key component of data governance. It provides a basis for understanding and agreement across divisions on the goals, benefits and activities necessary to manage data across the enterprise. The **EDMS describes the steps necessary to define the data management processes, techniques, and products required to achieve optimal sharing of Medicaid enterprise information.** Like other states, DCH originally built Medicaid systems to satisfy its own and federal business requirements, but did not design them to interoperate seamlessly with other systems within the enterprise, let alone with external systems and applications.

To achieve increased levels of MITA maturity, DCH needs to extend its current data and information activities to include data sharing, seamless integration, reuse, and semantic interoperability at the enterprise level, while maintaining data quality and integrity.

The EDMS is a guide to coordinate the activities of managing data, with the goal of getting the right data to the right people at the right time. In the future, DCH may decide to replace the traditional way of using interfaces to address many of its data sharing needs by developing standardized message formats containing the same data elements and formats to exchange data with external state, regional, and national entities.

The EDMS will be used to facilitate and develop information effectively shared across internal and external boundaries to improve mission performance. Implementation of the EDMS will provide the techniques, processes, and products to meet the need for timely and accurate information. It will also provide an impetus for DCH to better understand its data and how it fits into the total pool of Medicaid data. It addresses fundamental aspects (e.g., syntax and semantic interoperability) to enable information-sharing opportunities and to position DCH to operate in an environment of global information.

As stated in the MITA IA Information Capability Matrix, to reach MITA maturity level 3, an enterprise data management must include these capabilities:

- **Adoption of governance process and structure;**
- **Adoption of intrastate metadata repository;**
- **Adoption of intrastate enterprise modeling; and**
- **Adoption of statewide standards for data definitions, data semantics and harmonization strategies**

Within the Supporting Information section of this deliverable, the EDMS Roadmap shows the steps necessary to adopt these tools and processes in order to fully implement the DCH MITA Enterprise Data Management Strategy.

Scope

CMS defines the scope of the MITA DMS as addressing common Medicaid data management strategies, techniques, and components at a high level. CMS requires all individual state Medicaid agencies (SMAs) such as DCH to extend the DMS to include its unique strategies, techniques, and components, including specific technical systems components associated with existing physical data models, databases, and data files.

The following sections extend the CMS DMS to strategies specific to DCH and its Medicaid information management.

Information in Scope for DCH Medicaid Enterprise Data Management

The scope of DCH Medicaid data management includes the data used by its Medicaid business processes and programs, as will be shown in the DCH Enterprise Conceptual Data Model (ECDM). The nine areas depicted in the diagram below are based upon the MITA defined business areas with the Eligibility and Enrollment area combined with both Member and Provider data.

Figure 2: Data Areas in Scope for DCH MITA Enterprise Data Management

Audience

The EDMS is intended for the Information Architecture leader and the Data Administration (DA) Team. The DA Team will help to implement the DCH Enterprise Data Management Strategy (EDMS) and Enterprise Data Standards and Management Plan (EDSMP), facilitating effective sharing of information across internal and external boundaries to improve mission performance. EDMS implementation will add techniques, processes, and products to meet the need for timely and accurate information.

The DA Team will deploy the techniques and processes outlined in the EDMS and develop and maintain work products to support ongoing data management. These processes and products include metadata documentation, enterprise data analysis and modeling, training, and enforcement of standards across systems.

The DA Team will aid in the efforts to reach greater MITA maturity. The role of team members is to focus on data architecture at an enterprise business level, with a focus on data quality and data integration for systems interoperability.

Benefits of Enterprise (Medicaid) Data Management

As stated by CMS, the complexity and stovepipe nature of the current implementation of state systems have resulted in a lack of interoperability and poor exchange of data. Enterprise data management aids in information interoperability and exchange, providing strategies to combine tools, procedures, and processes to handle future data needs and achieve the following benefits:

- Aligns enterprise information-related activities and provides a roadmap to use in planning.
- Provides guidance for making decisions associated with information, data sharing, and seamless interoperability.
- Reduces cost by aligning and focusing information-related activities, e.g., identifying utilization anomalies, establishing, managing, and tracking compliance incidents.
- Provides the foundation to obtain data about beneficiaries, providers, and procedures and combined with other data, find billing aberrancies or outliers to support program integrity, fraud prevention, and detection.
- Provides an information structure that enables systems to share data in formats with a common definition resulting in consistent application across the enterprise.
- Reduces risk-to-system development by reducing custom solutions and promoting interoperability and data sharing.
- Increases overall quality of how data is managed and how information is exchanged.
- Provides a common set of processes, tools, and solutions for the information needs of Medicaid.
- An increased understanding of the enterprise architecture and data through education and effective uses of data, while ensuring data is visible, accessible, and understandable.

Increasing MITA Maturity through Data Management across the Enterprise

MITA maturity is increased through data governance and enterprise data management by the following activities:

- Expanding the capability to share data and improving universal data sharing through a standardized MITA Business Architecture, Information Architecture and Technical Architecture Framework to integrate both structured and unstructured data.
- Adopting an enterprise-wide view of data and supporting organizational roles in order to promote and ensure trusted data models to support all business areas and control redundancy.

Data Governance and its Role in Standards Adherence

Data governance is the overarching process in which an Enterprise Data Management Strategy can be implemented. Roles and responsibilities for data governance and implementation of activities to implement the EDMS are shown in Table 1, Data Management Roles and Responsibilities found in the Supporting Information section.

As defined by CMS, Data Governance defines the governance processes for making enterprise-wide decisions regarding information holdings. It provides the capability to determine ownership and data standard adoption processes, to address data integrity, to define processes for business-process development, and to establish a mechanism for arbitrating differences. The benefits are that it decreases data duplication, improves cost effectiveness of data sharing throughout the enterprise, and increases data quality. Adhering to enterprise data standards is a key aspect of data governance.

A new Data Administration Team to maintain the architecture of enterprise data from a conceptual and logical view is recommended, to report up to the MITA Executive Governance Committee.

Examples of Data Governance Activities

CMS provides the below listed examples of Data Governance EDM activities[2]:

- Registration of data solutions into an enterprise repository.
- Design of all proposed data solutions by providing logical data models and following data standards before passing the solution to the Database Administrator (DBA) for physical data modeling.
- Approval of all logical data models (for proposed data solutions, including those to be achieved by vendors).
- Enforcement of compliance standards for all data solutions with enterprise data-naming standards.
- Enforcement of compliance standards for all data solutions with security and disclosure.
- Data integration plan that supports a Technical Service Model (TSM), as part of an Enterprise Technical Management Strategy as recommended by the MITA Technical Architecture Framework, efforts for all data solutions.

[2] CMS MITA v.3, Part II- INFORMATION ARCHITECTURE, Chapter 2 – Data Management Strategy.

Standards Adherence Fostered by Enterprise Data Management

Enterprise data in conceptual and logical data models will use vocabulary standards (for data class names) preferred by national standards organizations, as follows:

- National Information Exchange Model (NIEM) version 3, for “core data” class names such as ORGANIZATION and PERSON, with ENTITY as the superclass for any business party; ACTIVITY to represent a business event; IDENTITY; LOCATION and ADDRESS; and use of “ASSOCIATION” naming for intersection of many-to-many relationships among classes;
- Health Level 7 (HL7) Reference Information Model ; and
- Applicable HIPAA Transaction Code sets.

DCH Information Asset Management

DCH should adopt a formal Information Asset Management protocol. Each information asset for which the state entity has ownership responsibility should be inventoried and identified to include the following:

- Description and value of the information asset.
- Owner of the information asset.
- Custodians of the information asset.
- Users of the information asset.
- Classification of information.
- FIPS Publication 199 categorization and level of protection (Low, Moderate, or High).
- Importance of information asset to the execution of the state entity's mission and program function.
- Potential consequences and impacts if confidentiality, integrity and availability of the information asset were compromised.

Implementation of the EDMS will enable DCH to better address each of these items, involving the organizational roles shown in Table 1, Data Management Roles and Responsibilities which is located in the Supporting Information section

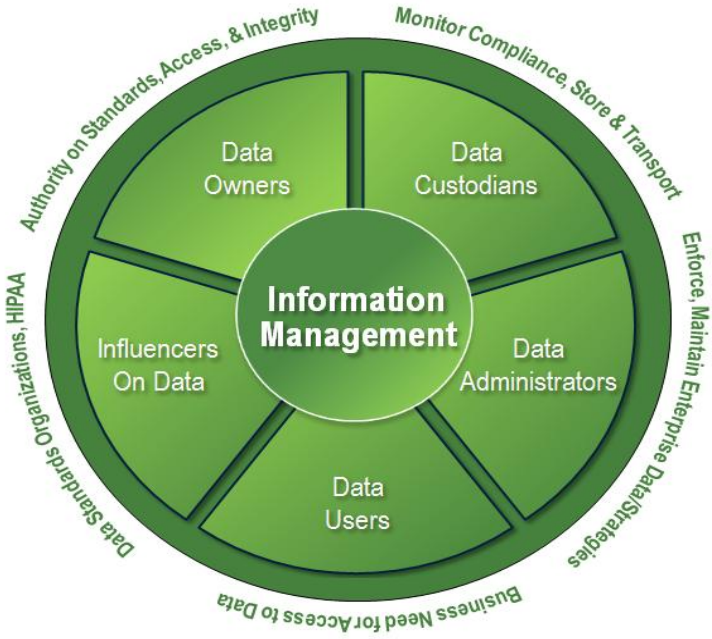

Figure 3: Dimensions of Information Asset Management

Components of the EDMS

Implementing the EDMS includes addressing the **business flow of data across DCH**, involving architecture, modeling, standards, metadata, management, interoperability, security and privacy, access methods, quality, and performance standards. These areas are described in the following sections.

Common Data Architecture

Common Data Architecture is the term used to describe the establishment of standard data management procedures for the data models and guidelines for data documentation, data-sharing development and use applicable to both structured and unstructured data, and management of metadata of all types. These guidelines ensure that all parties to system design/redesign efforts utilize pre-defined data entities and attributes, data models, and relationships to convey the overall meaning and use of Medicaid data and information. **The Seven Standards and Conditions, State Self-Assessment (SS-A), MITA Maturity Models (MMM), Concept of Operations (COO), and Business Process Models (BPM) provide a foundation for common data architecture.** The goals of common data architecture include:

- Defined data to support the business "To-Be" goals.
- Improved consistency in the development and deployment of systems.
- Increased the data quality and performance of systems.
- Normalization of data across applications.
- Decreased complexity of the Extract, Transform, and Load (ETL) process
- Decreased resource expenditures involved in new or redesigned system efforts.
- Identification, documentation, and limitation of the number of source data locations and access points.
- Reduced redundant data by improving communication about available classes and attributes, as well as assigning owners to promote modeling responsibility.

Enterprise Modeling of Data

With the MITA project, DCH is embarking on the modeling of enterprise data, to provide concepts and structures for department-wide system data elements that will more successfully interoperate.[3] Modeling of data provides diagrams to graphically depict the different data elements used by a program area or by the enterprise as a whole. Enterprise data modeling standardizes data across data source systems and third-party resources. Its benefits are increased model effectiveness, increased data sharing, decreased resource expenditures, and increased enterprise knowledge. It also helps move DCH toward consistent enterprise data standards through the adoption of common standards for data modeling policy, naming, classes, attributes, and data sets.

The Seven Standards and Conditions, State Self-Assessment (SS-A), MITA Maturity Model (MMM), Concept of Operations (COO), and Business Process Models (BPMs) provide a foundation for enterprise

[3] See also: DCH MITA Conceptual Data Model Deliverable Expectation Document (DED).

modeling, together with the Enterprise CDM and LDM to be developed on the MITA project[4], which show the scope and structure of MITA enterprise data across DCH.

Enterprise Metadata Repository

To aid the understanding of enterprise data held in systems and the definitions and formats of that data (the "metadata"), an enterprise metadata repository is required to enable access by all stakeholders to enterprise metadata. This allows for future systems to be designed and shared in a manner promoting data standards and interoperability. Future systems designed with reusable enterprise data structures and metadata will be more easily shared.

CMS states that the benefits to using an enterprise metadata repository include increased ability to obtain data at the right time, increased data quality, and decreased resources required to maintain data. State solutions should promote sharing, leverage, and reuse of Medicaid artifacts, technologies, and systems within and among States. Note that CMS itself plans to release more information as repositories and registries are established at the federal level.

For GA DCH, the metadata repository is currently within the Erwin tool.

Data Standards[5]

The EDMS provides mechanisms[6] to monitor and influence the environment in which Medicaid operates with respect to both the data landscape and the enabling technologies to achieving enterprise data sharing and interoperability. The EDMS Roadmap, found in the final section of Supporting Documents, outlines the roles, responsibilities, and tasks needed to achieve these goals. Data landscape refers to the broad range of national initiatives, standards organizations, and other organizations engaged in defining or influencing standards such as the Council for Affordable Quality Healthcare (CAQH) and the National Medicaid EDI Healthcare (NMEH) workgroups.

Enabling technologies refer to open standards, protocols, middleware, and other mature or emerging technologies that facilitate data sharing. The EDMS can be used to prioritize activities based on business needs and measures the resulting business value, for example, by choosing to first focus on common data involved in disparate systems across the enterprise. Business value will continually evolve as standards and data-sharing solutions become more refined. Refining standards does not necessarily mean increasing the number of standards; rather, standards apply where they provide evolving business value.

CMS has published a list of standards organizations that are relevant to the MITA framework:

- **Development Standards and Maintenance Organizations:** Accredited Standards Committee (ASC) Standard X12N, HL7, and National Council for Prescription Drug Programs (NCPDP)

[4] Ibid.
[5] See also: DCH MITA Team's Data Standards and Management Plan (to be developed).
[6] Such as: Data administration as a new role type; a metadata repository for shared access to enterprise data definitions, structures and standards; use of subsets of enterprise conceptual and logical data models for each program area, and the data standards to which enterprise data must adhere.

- **Standards, Consortiums, Oversight and Advisory Organizations**: Office of the National Coordinator for Health Information Technology (ONC), American National Standards Institute (ANSI), and National Health Information Network (NHIN)
- **Vocabulary Domains and Data Sets:** American Medical Association's Current Procedural Terminology (CPT)-4, Diagnosis Related Group (DRG), International Classification of Disease (ICD)-9, ICD-10)
- **Communication Information Structures:** Unified Modeling Language (UML) V2/Unified Modeling Methodology (UMM), X12N)
- **Other Federal Agency Projects or Programs** (NIEM, Federal Health Information Model (FHIM))

CMS indicates that many of these standards are critical for the common exchange and interpretation of data within the SMA and that applicable standards must be identified and adhered to with the publishing of the conceptual and logical data models. The models must support the applicable standards, and it is recommended that, once developed, the Data Administration Team (or other group as designated by DCH) be assigned to governing adherence to these standards.

Vocabulary Data Standards

Vocabulary standards will be defined where possible using NIEM, which began as the result of a collaborative effort by the United States Department of Justice and Department of Homeland Security to produce a set of common, well-defined data elements to be used for data exchange development and harmonization. Naming for applicable data classes will follow predefined names preferred by NIEM and HL7:

- NIEM version 3, for “core data” class names such as ORGANIZATION and PERSON, with ENTITY as the superclass for any business party; ACTIVITY to represent a business event; IDENTITY; LOCATION and ADDRESS
- Health Level 7 Reference Information Model for health-care specific data classes

Data Element Format Standards for Enterprise Data (Conceptual and Logical)

NIEM does not have data classes or elements defined for some of the health-care specific data classes necessary for an SMA data model. For these and other areas that are not supported by NIEM conventions, the data naming and formats will be supplemented by using standards from one of the other standards organizations, such as HL7 or ASC X12N.

- Applicable data naming and data elements from NIEM Core and other NIEM datasets
- Applicable data naming and data elements from HL7 datasets

Data Element Format Standards for Physical Data Sets (Existing Systems/Transactions)

Other federal agencies are in the process of transforming their models to NIEM. According to the CMS MITA v3, Part II IA Chapter 5 – DATA STANDARDS, the FHIM team is writing rules to transform models to NIEM. FHIM is a national program supported by the federal government that provides a community of users, tools, common terminology, governance, methodologies, and support that enables enterprise-wide information exchange. Once this is complete, it is the intent of CMS to design and develop an information exchange model and tools that are fully compliant with NIEM requirements as part of the Health Insurance Exchange (HIX), Medicaid and Children’s Health Insurance Program (CHIP) operations. MITA

supports the use of the related FHIM development framework as the method that MITA users can demonstrate conformance to NIEM and develop interfaces.

- Applicable HIPAA Transaction Code sets, e.g.: CPT-4, ICD-10, and code sets from Health Care Common Procedure Coding System (HCPCS), and National Drug Code (NDC), etc.

Data-Sharing Architecture

Data-Sharing Architecture is the term used by CMS to describe the technology considerations for DCH program and technical staff (as well as vendor staff, as appropriate) to share enterprise information and standards for emerging data solutions (new or redesigned). Based on business requirements and MITA considerations, CMS recommends that Data Administration staff be responsible for the following aspects of data-sharing architecture:

- Defining and maintaining data and information exchange formats for future system efforts.
- Adopting MITA-standard data definitions and data-sharing schemas where possible.
- Maintaining a centralized meta-data repository[7] of this information for general use.

For CMS, data-sharing architecture (including the meta-data repository and associated modeling toolset) also addresses data semantics, data harmonization strategies, shared-data ownership, security and privacy implications of shared data, and the quality of shared data. It is a key technology to enable state solutions to promote sharing, leverage, and reuse of Medicaid technologies and systems within and among States, thereby reducing costs and achieving further levels of MITA maturity.

The EDMS identifies enabling technologies and specifies data design structures to enable data interoperability, along with associated processes and procedures. The resulting IA leads to the development of the target or To-Be data management environment. New projects involving data architecture within DCH and where practical, changes to existing architectures, must adhere to the enterprise standards, which will be fully documented in the MITA Team's Data Standards and Management Plan.

This means that new systems will need to map the physical data source to the logical data model to ensure that data can be shared through well-defined interfaces. For domain standards, future systems must use the same coding standards and data sets to maintain context across enterprise systems. Enforcement of these standards is addressed in the discussion of roles within data management. Following the definitions defined by these vocabulary domain and datasets, groups will achieve this contextual integrity.

[7] DCH currently uses ERwin for data modeling, and could use ERwin as a meta-data repository as well

Requirements for Enterprise Data Sharing Architecture Tools

Enterprise data sharing architecture requires the following features and abilities:

Enterprise Data Repository Capabilities

- Ability to share and reuse (via cross-references and cascading inheritance) documented objects across the development life cycle, from analysis of requirements to system maintenance.
- Ability to support documentation for enterprise data elements, to include at a minimum, documentation of meta data for:
  - data classes (entities), data sub-classes and super-classes (subtypes and super-types)
  - data attributes
  - data element definitions at both the class and attribute level
  - relationships between classes
  - keys or unique identifiers among attribute and relationship sets
  - program area entity and attribute cross-references (in a many-to-many configuration)
  - glossary definitions for both data classes and data attributes, exportable to report formats
- Ability to matrix enterprise data concepts to business areas as subset data concepts as well as to lower levels in the system development lifecycle (conceptual data class to logical data class, to physical schema element).
- Ability to matrix data elements to business processes and system components in both Create-Retrieve-Update-Delete (CRUD) matrix form and matrices to cross-relate data at rest to data in motion (XML schema packages, system functional components).
- Ability to restrict views of repository meta-data according to viewer role type involving specific access privileges (read-only, create, update, delete).

Graphical Diagramming Capabilities

- Diagramming capacity to include subsets and entire enterprise level of data elements; color-coding ability of classes for diagrams, etc.
- Ability to graphically model, document and separately define:
  - data classes (entities)
  - data sub-classes and super-classes (subtypes and super-types)
  - data attributes
  - data element diagram legend, title and date
  - relationships between classes
  - keys or unique identifiers among attribute and relationship sets
- Ability to model data and data interrelationships using UML as a standard diagramming notation, with an option for entity-relationship diagramming.
- Ability to dynamically link (double-click) data classes on diagrams to underlying repository meta-data components for the class, its attribute detail, etc.
- Ability to restrict views of diagrams according to viewer role type involving specific access privileges (read-only, create, update, delete).

Configuration and Licensing Capabilities

- Ability to provide an "alert" notification of a diagram viewer that underlying elements may have changed, requiring diagram update if synchronization is suitable.
- Ability for the repository and diagrams to be accessed in (at least) read-only viewing modes across the intranet in order to share models and meta-data on MITA enterprise data statewide (MITA Level 3 maturity item), without requiring user licenses for read-only access to tool outputs.
- Ability for repository subsets and diagrams to be configured using version control.
- Ability to define attribute data standards and to "inherit" these standards among logical and physical layers of data elements for enforcement of standards.
- Ability to transform a conceptual data model to logical and physical data elements, to support the full life cycle of systems development from strategic planning and enterprise IA to system design and maintenance.
- Ability to generate C# and XML packages against data class/attribute sets.

Security and Privacy of Enterprise Data Assets

The groups involved in data governance, including the MITA Executive Governance Committee or their delegates, will develop and monitor performance standards for the maintaining the security and ensuring the privacy of enterprise data assets.

Information Quality and Performance Standards

The groups involved in data governance, including the MITA Executive Governance Team or their delegates, will develop and monitor performance standards for information quality, accessibility and interoperability.

Using the EDMS

Using the EDMS is critical to the successful transformation and evolution of DCH MITA maturity, to help transition its current information architecture towards a MITA IA and the Seven Standards & Conditions.

Key Activities to Implement the EDMS

To maintain an effective and coherent EDMS, key activities that the MITA PMO will need to ensure are implemented include:

- Collaboration – Participating in management discussions to review and receive input from stakeholder management on desired or expected EDMS goals, activities, and areas of concern
- Definition – Defining of parameters and scope of the EDMS on an ongoing and updated basis
- Using the EDMS along with the Business Capability Matrix (BCM), the Information Architecture Capability Matrix (ICM), and the Technical Architecture Capability Matrix (TCM) in planning future interoperability of data solutions
- Ensuring that processes to support the EDMS appear in solicitation documents (Requests for Offer (RFO) and Requests for Proposals (RFP)) and requiring that adherence to MITA enterprise data standards and definitions to be part of the evaluation criteria in vendor responses (and part of the performance measures on which vendor projects must be measured)

Creating a New Data Administration Function

The EDMS will be facilitated in its implementation by creation of a new Data Administration function within DCH that reports up to the MITA Executive Governance.
CMS defines the Data Administration role to involve:[8]

- Guiding the creation and monitoring the usage of data and information as vital agency assets
- Promulgating agency standards, procedures and guidelines related to data names and definitions
- Maintaining the inventory of enterprise-wide data assets
- Facilitating understanding of the meaning, accuracy and timeliness of data assets
- Promoting the reuse of standardized data names, definitions, elements and values

Integrating EDMS with Existing Systems

The new Data Administration Team will have responsibilities to extend the EDMS in mapping to existing systems:

[8] From CMS.gov - http://www.cms.gov/Research-Statistics-Data-and-Systems/CMS-Information-Technology/DataAdmin/index.html

- Extend the EDMS associated with the physical datasets and databases of existing data solutions, by cross-referencing CDM and LDM data elements and meta-data to existing systems and mapping of data from existing systems to logical data model elements
- Implement the strategy for integrating and harmonizing physical data (in existing systems) into an Operational Data Store (ODS) of actual data content, using the logical data model to define the to-be data classes and attributes within the ODS. The process of data harmonization involves the analysis of data from multiple sources, determining the commonality between current systems' data elements prior to merging the common data into the matching ODS data tables and fields
- Require data administrative approval prior to altering common data elements in DCH systems

Note that these activities are critical to Master Data Management which is discussed in detail in the following section.

Implementing Master Data Management

Master Data Management (MDM) manages the ***actual data content*** of systems, mapping and matching system elements to "master" data to be used or inherited for use across the enterprise. MDM comprises the processes, governance, policies, standards and tools that consistently define and manage the critical data of an organization to provide a single point of reference.[9] An established EDMS, and the development of enterprise data models with associated "master data" definitions, is a prerequisite for MDM.

MDM activities include source identification, data collection, data transformation, normalization, rule administration, error detection and correction, data consolidation, data storage, data distribution, data classification, taxonomy services, item master creation, schema mapping, product codification, data enrichment and data governance (of actual data held within systems). In the course of transformation and normalization, data administrators adapt descriptions to conform to standard formats and data domains, making it possible to remove duplicate instances of any entity. Such processes generally result in the enterprise data repository, from which all requests for a certain entity instance produce the same description, irrespective of the originating sources and requesting destination.

MDM involves collecting, aggregating, matching, consolidating, quality-assuring, persisting and distributing data throughout an organization to ensure consistency and control in the ongoing maintenance and application use of this information. The result is a more clear distinction between three types of data used by an organization, and the ability to map or inherit from the source "master" data:

- Master data – the unique business data used in transactions, and dimensions for analysis
- Transactional data – the data that supports applications
- Analytical data - supports decision making

[9] "What is Master Data" SearchDataManagement, TechTarget, 22 November 2010, http://searchdatamanagement.techtarget.com/definition/master-data-management

How is data transformed by MDM?

Master data management of disparate data systems requires data transformations; the data extracted from the disparate source data system is transformed and loaded into the master data management hub. To synchronize the disparate source master data, the managed master data extracted from the master data management hub is again transformed and loaded into the disparate source data system as the master data is updated.

MDM Tools

Tools exist to merge, match and transform data to enable MDM. Such tools create an authoritative source of master data by examining existing systems and transforming the data through activities such as:

- Standardizing data
- Removing duplicates
- Incorporating rules to eliminate incorrect data from entering the system

MDM tools can include data networks, file systems, a data warehouse, data marts, an operational data store, data mining, data analysis, data virtualization, data federation and data visualization. One of the newest tools, virtual master data management, utilizes data virtualization and a persistent metadata server to implement a multi-level automated MDM hierarchy.

Challenges of MDM

Problems in achieving MDM include issues with the quality of data, consistent classification and identification of data, and data reconciliation. As with other Extract, Transform, Load-based data movements, processes to achieve true MDM are expensive and inefficient to develop and to maintain which greatly reduces the return on investment for MDM. Reconciling several master data systems can present difficulties because of the dependencies that existing applications have on the master databases. As a result, more often than not the two systems do not fully merge, but remain separate, with a special reconciliation process defined that ensures consistency between the data stored in the two systems. Reconciliation processes can become extremely complex, and consequently unmanageable and unreliable, causing serious operational problems in the areas of customer satisfaction, operational efficiency, decision-support, and regulatory compliance.

The adoption and implementation of the EDMS will lay the foundation for DCH to further explore the benefits of MDM to the Medicaid Enterprise in the future. Detailed roles and responsibilities are provided in Table 1, Data Management Roles and Responsibilities. Proposed timelines to achieve these actions is shown in the Roadmap section below in Table 2.

Supporting Information

The EDMS has described activities needed to support the implementation of an overall data management program. Figure 4 below is one way to group the processes involved in enterprise data management.

Strategic Planning	Data Governance	Data Architecture	Data Warehousing	Data Quality	Data Security
Perform Enterprise Data Management Strategic Planning	Define Data Governance Process	Develop and Maintain Enterprise Data Models	Develop and Maintain Enterprise Dimensional Data	Develop and Promote Data Quality Best	Develop and Implement Data Security Standards
Define Data Policies	Implement Data Governance Process	Develop and Maintain Modeling and Design	Develop Data Warehouse and BI Architecture	Define Data Profiling Process	Develop and Implement Data Privacy Standards
Promote Compliance with Data Policies & Standards	Create, Capture and Maintain Enterprise Metadata	Establish and Maintain Enterprise Data Architecture	Develop Data Warehouse and BI Technical Standards	Manage Data Quality	Develop and Promote Data Security & Privacy Best Practices
Publicize and Promote Data Management	Develop and Implement Enterprise Metadata Architecture	Develop and Maintain Data Integration and		Measure Data Quality	Perform Data Security Auditing
Identify and Justify Resources and Budget Needs	Create and Maintain Master Data Management (MDM) Standards	Provide Support to Database Administrators (DBAs)			
Monitor EDM Program Performance		Assess Application Integration Interfaces			
Coordinate with External Standards Organizations		Evaluate Tools			

Figure 4: Enterprise Data Management and Related Processes

Organizational Management Roles and Responsibilities for EDMS

Table 1 describes data management roles and responsibilities.

Data Management Roles and Responsibilities	
Data Owners	Data Owners authorize access to information in accordance with the classification of the information and the need for access to it. Data Owners monitor and ensure compliance with all applicable laws, security policies and procedures affecting the information. Data owners create, manipulate and report on data from within a business program perspective. To achieve the goals of enterprise data management, data owners support information system projects in the following ways: ▪ Provide data requirements, characteristics, formats, and definitions for program area data. ▪ Provide system data outputs, interfaces and plans for redesign or enhancement of system data.
Data Users	Data users comply with applicable laws and administrative policies as well as any additional security policies and procedures established by the owner of the information and the agency Information Security Officer. Data users are users of the system data, and may also create, manipulate and report on data required from a program role perspective.
Data Custodians* ***Data Owners control access; Data Custodians control security.**	Data Custodians monitor and ensure compliance with all applicable laws, and agency and state security policies and procedures affecting the information, advising data owners and the agency ISO of vulnerabilities, specific means of protecting that information, and any actual or attempted violations of security policies, practices and procedures.

Data Management Roles and Responsibilities	
Data Administrators ****Control data structure and interoperability.**	The data administrator role is four-fold, as shown below, managing conceptual data structures, with a focus on meta-data used across the enterprise and information systems. They support information system projects throughout the enterprise through the following activities:
	Support Development of Enterprise Data Management Strategies and Standards ▪ Participate in the development of data governance processes and procedures with the proposed new MITA Executive Governance Committee to promulgate DCH enterprise-wide understanding of master data management and common data architecture principles and policies. ▪ Define and maintain Common Data Architecture (CDA) standards and procedures. ▪ Solicit and follow-up on feedback from projects and management to improve MITA enterprise data standards and procedures.
	Ongoing Maintenance of Enterprise Data according to Standards ▪ Administer and maintain the DCH MITA Enterprise Metadata Repository containing approved sets of enterprise data vocabulary/names, definitions and formats. ▪ Create and maintain naming conventions based on approved MITA enterprise data standards such as NIEM and HL7. ▪ Maintain an Enterprise CDM containing common and reusable data objects at a high level (“concepts”) for use as standardized data templates for new software development projects or legacy re-engineering. ▪ Maintain an Enterprise Logical Data Model (ELDM) which derives its conceptual structure from the ECDM, adding detail; the ELDM contains common and reusable data objects at a fully-attributed level as well as standardized data modeling templates for jump-starting new software development projects or re-engineering legacy applications. ▪ Maintain program-specific CDMs and LDMs using enterprise data concepts and standards. ▪ Map and maintain linkages to enterprise data to existing system elements.

Data Management Roles and Responsibilities	
	Promulgation of Enterprise Data Management Strategies and Standards ▪ Publish standardized data naming and format conventions in the enterprise repository for data standards enforcement. ▪ Communicate to the enterprise what shared data elements and standards exist and where they can be found via the intranet access to the repository (or its outputs), together with any updates as required, on a periodic basis. ▪ Provide training and outreach to business SMEs, analysts, consultants and IT developers on the use of enterprise data and standards as appropriate.
	Enforcement of Enterprise Data Management Strategies and Standards These responsibilities are key to promoting master data management across DCH. ▪ Ensure that data owners and data custodians/stewards approve any changes to the program-specific information used in business process CDMs and LDMs. ▪ Assist in ensuring that DCH solicitations for IT procurements and any vendor responses refer to and comply with the required use of enterprise data model structures, vocabulary, definitions and formats. ▪ Review and approve project-specific adherence to CDM/LDM standards, and associated vocabulary and data element formats, on both vendor-managed and State-managed system development projects. ▪ Demonstrate source-to-target data lineage among systems and objects. ▪ Identify the systems and business processes that do not use data standards and consider which data conversions are necessary. ▪ Perform data analysis and impact analysis services in conjunction with new and ongoing application development efforts.
Data Influencers	Many business parties influence how data is structured, shared and protected. At DCH, these may include: ▪ Georgia Health Information Technology Regional Extension Center (GA-HITEC) ▪ Operations ▪ Finance ▪ Office of Inspector General ▪ Office of Information Technology

Data Management Roles and Responsibilities	
MITA Executive Governance Committee	MITA Governance at DCH has the following objectives: ▪ Apply an enterprise-wide approach to ensure that DCH advances in MITA maturity. ▪ Monitor implementation of DCH "To-Be" goals and commitments. ▪ Review and approve all Medicaid Management Information System (MMIS)-related concepts and projects, ensuring alignment with the DCH Strategic Plan and MITA State Self-Assessment (SS-A) goals. ▪ Communicate MITA-related goals, objectives, and status Department-wide. ▪ Manage the assignment of DCH owners for MITA Business Processes. ▪ Review and approve all requests for enhanced federal funding. ▪ Review and approve all DCH budget requests, ensuring alignment to the DCH Strategic Plan and MITA SS-A goals. ▪ Sponsor the creation of an enterprise services infrastructure. ▪ Ensure continued support and buy-in of MITA governance from DCH executive staff. ▪ Develop and monitor performance standards for information quality, accessibility and interoperability

Data Management Roles and Responsibilities	
Health Information Management Committee (Recommended)	We recommend the development of a Health Information Management Committee to provide overall governance for data-related issues. Its role includes: ▪ Prioritizes and aligns resources to address critical programmatic needs. ▪ Obtains funding to support technology development. ▪ Identifies Data Owners and their roles and responsibilities. ▪ Reviews data delivery processes to assess ability of shared departmental resources to meet programmatic needs, and develops specific criteria for prioritizing health information projects, and specific processes for ranking those projects. The processes it uses to accomplish this mission include: ▪ Collaboration – Participating in management discussions to review and receive input from stakeholder management on desired or expected EDMS goals, activities, and areas of concern ▪ Definition – Defining of parameters and scope of the EDMS on an ongoing and updated basis ▪ Using the EDMS along with the Business Capability Matrix (BCM), the Information Architecture Capability Matrix (ICM), and the Technical Architecture Capability Matrix (TCM) in planning future interoperability of data solutions ▪ Ensuring that processes to support the EDMS appear in solicitation documents
Data and Research Committee (DRC) (Recommended)	We recommend that the MITA Executive Governance Committee develop a Data and Research Committee (DRC). The DRC oversees DCH's data request evaluation process from external entities requesting protected data from DCH for research and public health purposes. In this context, protected data refers to any data containing confidential information that otherwise would be unavailable to the public. The DRC assesses the appropriateness of requests for protected data, assigns a priority status to each request, and recommends potential approval/denial action to DCH executive management.

Table 1 – EDMS Roles and Responsibilities

ROADMAP for EDMS Implementation

Initial steps and responsibilities to implement the DCH EDMS are identified in the Roadmap below. Timelines for these activities are yet to be determined.

ROADMAP ACTIVITY and TIMELINE FOR ENTERPRISE DATA MANAGEMENT STRATEGY			
1.	Acquire necessary licenses for a repository/modeling toolset to support enterprise-wide information architecture	PMO	TBD
2.	Ensure procurement process and Statements of Work include requirements for adherence to EDMS standards, including data naming and format characteristics, for system design/redesigns	PMO	TBD
3.	Collaborate across stakeholders to develop and approve a charter for a new Data Administration (DA) Team to manage enterprise data	MITA Executive Governance Committee	TBD
4.	Identify data administrators and data analysts within the organization, capable of assuming new responsibilities based on the DA Team charter	MITA Executive Governance Committee	TBD
5.	Clearly identify data owners and recognize their roles and responsibilities; requires education first, and adoption of EDMS recommendations	MITA Executive Governance Committee	TBD
6.	Establish/communicate business-area data standards: ▪ Vocabulary data standards (class and attribute names) ▪ Domain entity vocabulary (subtype classes of enterprise data) ▪ Structure data standards	MITA Executive Governance Committee or delegates	TBD
7.	Implement enterprise data management (IA) repository	Data Administration Team	TBD
8.	Establish workgroups of data administrators and stewards to approve CDMs and LDMs on system development/redesigns	MITA Executive Governance Team	TBD
9.	Maintain and publish Enterprise CDM content using the new repository toolset, for common data model elements across the enterprise (based on the ECDM)	Data Administration Team	TBD
10.	Create, maintain and publish CDM subsets by business area utilizing common data model elements (based on CDM subsets)	Data Administration Team	TBD
11.	Create, maintain and publish ELDM content for fully-attributed common data model elements (based on ELDM subsets)	Data Administration Team	TBD
12.	Create, maintain and publish LDM subsets by business area utilizing common data model elements (based on LDM subsets	Data Administration Team	TBD

Table 2 - EDMS Implementation Roadmap: Activity Timeline

Appendix 1: Supporting Documents

Documents providing input to the data model deliverables described above include:

- CMS Data Administration Operating Procedures Guidelines: http://www.cms.gov/Research-Statistics-Data-and-Systems/CMS-Information-Technology/DataAdmin/Downloads/OperatingProceduresGuidelines20131201.pdf
- CMS MITA Framework 3.0 Part II- Information Architecture:
  - Chapter 2, Data Management Strategy
  - Chapter 5, Data Standards
  - Chapter 6, Information Capability Matrix
- Federal Enterprise Architecture Program, Federal Data Reference Model: http://www.whitehouse.gov/sites/default/files/omb/assets/egov_docs/DRM_2_0_Final.pdf
- HL7 International data standards: http://www.hl7.org/implement/standards/rim.cfm
- HL7 Fast Health Interoperability Resources (FHIR) Standards: http://hl7.org/implement/standards/fhir/index.html
- NIEM version 3.0 data for core systems: https://www.niem.gov/technical/Pages/current-release.aspx
- NIEM Naming and Design Rules: http://reference.niem.gov/niem/specification/naming-and-design-rules/1.3/niem-ndr-1.3.pdf
- Georgia DCH Strategic Plan: http://dch.georgia.gov/sites/dch.georgia.gov/files/related_files/site_page/DCH_Strategic_Plan_2013_2016_rv_1-3a.pdf

Appendix 2: CMS Operating Procedures for EDMS Implementation

CMS guidelines for Enterprise Data Management for MITA form one instance of best practice guidance, including CMS' own Data Management Operating Procedures and Guidelines[10] (updated December 1, 2013). CMS' DM procedures target development and reusability of data across different business areas and systems, and can be used by DCH to establish processes for the following data management tasks:

DM OP-002 OPERATING PROCEDURE FOR IDENTIFYING SYSTEM INTERFACES
DM OP-003: OPERATING PROCEDURE FOR DEVELOPING THE CONCEPTUAL DATA MODEL
DM OP-004: OPERATING PROCEDURE FOR ESTIMATING DATA MANAGEMENT SERVICE NEEDS
DM OP-005 OPERATING PROCEDURE FOR DEVELOPING THE LOGICAL DATA MODEL
DM OP-006 OPERATING PROCEDURE FOR REUSE OF ENTERPRISE ENTITIES, RELATIONSHIPS AND ATTRIBUTES
DM OP-007 OPERATING PROCEDURE FOR REUSE OF ENTERPRISE DATA RESOURCES
DM OP-008 OPERATING PROCEDURE FOR DEFINING DATA ENTITIES
DM OP-009 OPERATING PROCEDURE FOR NAMING DATA ENTITIES
DM OP-010 OPERATING PROCEDURE FOR DEFINING DATA ATTRIBUTES
DM OP-011 OPERATING PROCEDURE FOR ANALYZING TYPES OF DATA ATTRIBUTES
DM OP-012 OPERATING PROCEDURE FOR NAMING DATA ATTRIBUTES
DM OP-013 OPERATING PROCEDURE FOR MODELING DERIVED DATA-DELETED
DM OP-015 OPERATING PROCEDURE FOR DEFINING RELATIONSHIPS, CARDINALITY, AND OPTIONALITY
DM OP-016 OPERATING PROCEDURE FOR ASSIGNING A PRIMARY IDENTIFIER
DM OP-017 OPERATING PROCEDURE FOR NORMALIZING THE PROJECT LOGICAL DATA MODEL
DM OP-018 OPERATING PROCEDURE FOR DOCUMENTING DOMAIN VALUE RULES
DM OP-020 OPERATING PROCEDURE FOR DOCUMENTING DATA ISSUES
DM OP-021 OPERATING PROCEDURE FOR ASSIGNING INFORMATION SECURITY CATEGORIES
DM OP-022 OPERATING PROCEDURE FOR GENERATING THE PROJECT METADATA REPOSITORY
DM OP-026 OPERATING PROCEDURE FOR COMPLETING THE DATA MODELS
DM OP-027 OPERATING PROCEDURE FOR GRANTING PROJECT DATA MODEL LIBRARY ACCESS
DM OP-028 OPERATING PROCEDURE FOR NAMING AND DEFINING DATA MODELS
DM OP-029 OPERATING PROCEDURE FOR SELECTING MODEL TYPE
DM OP-031 OPERATING PROCEDURE FOR CAPTURING THE STANDARD LOGICAL DATA MODEL METADATA
DM OP-032 OPERATING PROCEDURE FOR CAPTURING THE STANDARD PHYSICAL DATA MODEL METADATA
DM OP-033: OPERATING PROCEDURE FOR RECORDING DATA MODEL CHANGES
DM OP-034: OPERATING PROCEDURE FOR UPDATING THE ELDM
DM OP-035: OPERATING PROCEDURE FOR REQUESTING A NEW STANDARD TERM.
DM OP-036 OPERATING PROCEDURE FOR ASSIGNING DATA ANALYSTS
DM OP-037 OPERATING PROCEDURE FOR CONDUCTING THE LOGICAL DATA DESIGN KICKOFF MEETING
DM OP-038 OPERATING PROCEDURES FOR DEFINING A SUBJECT AREA
DM OP-039 OPERATING PROCEDURE FOR CONDUCTING THE DATA ARCHITECTURE REVIEW
DM OP-040 OPERATING PROCEDURE FOR DESIGNATING REPRESENTATION TERM AND DATA TYPE
DM OP-041 OPERATING PROCEDURE FOR ASSIGNING DATE FORMATS
DM OP-042 OPERATING PROCEDURE FOR MODELING SUPER-TYPES AND SUBTYPES
DM OP-043 OPERATING PROCEDURE FOR MANAGING DATA VALUES THROUGH PHYSICAL CONSTRAINTS
DM OP-044 OPERATING PROCEDURE FOR PREPARING A PROJECT LOGICAL DATA MODEL FOR PHYSICAL DESIGN
DM OP-045 OPERATING PROCEDURE FOR CONSTRUCTING PHYSICAL TABLE OR FILE NAMES
DM OP-046 OPERATING PROCEDURE FOR CONSTRUCTING PHYSICAL COLUMN OR ELEMENT NAMES

[10] CMS Data Management Operating Procedures and Guidelines, December 1, 2013, at: http://www.cms.gov/Research-Statistics-Data-and-Systems/CMS-Information-Technology/DataAdmin/Downloads/OperatingProceduresGuidelines20131201.pdf

Appendix 3: Industry Best Practices

Other sources provided best practice implementation recommendations for the EDMS:

Defining the Centralized Data Ownership Policy

By David Loshin, Enterprise Knowledge Management, the Data Quality Approach, © 2001:

[On the Benefits of Central Data Ownership and Coordination]
In a centralized ownership model, there is a single entity (person or group) responsible for all data ownership for the entire enterprise. Centralization implies that all ownership activities are coordinated from a single point of control, as well as coordination of metadata, information sourcing, and so forth. Centralized ownership yields the benefit of the value added — and whether the costs associated with centralization are offset by it. The costs include the increased management overhead, bureaucracy, and system integration, among others. The benefits include enterprise standardization for data and systems, the ability to make use of merged data for additional knowledge discovery, and increased leverage when dealing with external data suppliers. [Emphasis added.]

Steps for defining a data ownership policy.

1. Identify the interested parties or stakeholders associated with the enterprise data. This includes identifying the senior level managers that will support the enforcement of the policy.
2. Catalog the data sets that are covered under the policy.
3. Determine the ownership models in place and whether these are to continue or will be replaced.
4. Determine the roles that are and are not in place. Assign the responsibilities to each role, and assign the roles to interested parties.
5. Maintain a registry that keeps track of policies, data ownership, roles, responsibilities, and other relevant information.

Cataloging Data Sets
Once the stakeholders have been identified, the next step is to learn what data sets should fall under the ownership policy. The stakeholders should be interviewed to register the data sets with which they are associated and the degree to which each believes his or her stake in the data is. The goal of this step is to create a meta-database of data sets to use in the enforcement of the data ownership policies. This catalog should contain the name of the data set, the location of the data set, and the list of stakeholders associated with the data set. Eventually, the catalog will also maintain information about data ownership and responsibilities for the data set.

Enterprise Data Management Roles and Responsibilities

David Loshin, Enterprise Knowledge Management, the Data Quality Approach, © 2001:

Identifying and Assigning Roles
The next step is to determine the roles that are associated with each set of data in the enterprise and describe the responsibilities of each role. Here are some examples, although this list is by no means meant to be exhaustive.

- Chief Information Officer - The CIO is the chief holder of accountability for enterprise information and is responsible for decisions regarding the acquisition, storage, and use of

data. He or she is the ultimate arbiter with respect to dispute resolution between areas of ownership and is the ultimate manager of the definition and enforcement of policies.

- Chief Knowledge Officer - The chief knowledge officer is responsible for managing the enterprise knowledge resource, which dictates and enforces the data sharing policies, as well as overseeing the general pooling of knowledge across the organization.
- Data Trustee - The data trustee manages information resources internal to the organization and manages relationships with data consumers and data suppliers, both internal and external.
- Policy Manager - The policy manager maintains the data ownership policy and negotiates any modifications or additions to the data ownership policy.
- Data Registrar - The data registrar is responsible for cataloging the data sets covered under the policy as well as the assignment of ownership, the definition of roles, and the determination of responsibilities and assignments of each role. The data registrar also maintains the data policy and notifies the policy manager if there are any required changes to the data ownership policy.
- Data Steward - The data steward manages all aspects of a subset of data with responsibility for integrity, accuracy, and privacy.
- Data Custodian - The data custodian manages access to data in accordance with access, security, and usage policies. He or she makes sure that no data consumer makes unauthorized use of accessed data.
- Data Administrator - The data administrator manages production database systems, including both the underlying hardware and the database software. The data administrator is responsible for all aspects related to the infrastructure needed for production availability of data.
- Security Administrator - The security administrator is responsible for the creation of and the enforcement of security and authentication policies and procedures.

  …
- Director of Application Development - The director of application development manages requirements analysis, implementation, testing, and deployment of new functionality for eventual turnover to the production facility.
- Data Consumer - A data consumer is an authorized user that has been granted access rights to some data within the enterprise.
- Data Provider - A data provider is an accepted supplier of information into the system.

These roles will then be integrated into a reporting structure where there are clear lines of responsibility corresponding to degrees of ownership. Note that some responsibilities are assigned to multiple roles, causing "role overlap," whose governance must be integrated into the reporting structure as well. At this point, the senior manager responsible for information (typically a chief information officer) will then assign ownership roles and responsibilities to the different organizational stakeholders.

The ownership registry is created from the data catalog and the assignment of roles. It is the enterprise log that can be queried to determine who has the ultimate responsibility for each data set. The ownership

registry should be accessible by all interested parties, especially when new data requirements arise or there is a conflict that needs resolution.

Enterprise Data Management Mistakes to Avoid

From David Marco, author of Universal Meta Data Models and Building and Managing the Meta Data Repository, in a speech given at the Enterprise Data World conference, as reported by Mark Brunelli, News Director, TechTarget.com, April 6, 2011:

Building silos
Enterprise information management programs, by definition, are enterprise-level initiatives that should touch every aspect of an organization, according to Marco. EIM programs may comprise several point or department-level projects designed to deliver incremental value over time, but each component of the program must be designed with enterprise-wide goals in mind. A silo can be defined as something that is disconnected from the rest of the enterprise. And if an organization is building an EIM program around just one application or just one line of business, then chances are they are building a silo.

"One thing that is always a concern is when you see an EIM initiative run as a standalone project," Marco said. "It's not supposed to be something that five people go off and try to hammer out. It's something that should have enterprise exposure at an enterprise level and, unfortunately, this doesn't happen as often as it should."

Failing to balance goals
Organizations taking on an EIM program need to balance long-term strategic goals with short-term tactical goals. This can be difficult because executives like to see results. One way to deliver incremental results, according to Marco, is to design the EIM program with this question in mind: What are some things that we can do in the next year that have tactical value but also move us down the organization's strategic path?

"Identify those optimum projects that tactically benefit the company but that also move the strategic long-term vision forward," Marco said. "And again, please don't make the mistake of building siloed short-term applications."

Boiling the ocean
Remember to take an incremental approach to EIM that delivers value over time. Telling company executives that the EIM program will deliver results in four years simply won't cut it, Marco said. Organizations need to deliver value earlier in the lifecycle. "I'm just a big believer in the iterative approach," Marco said. "I believe in three- to six-month project lifecycles."

Organizations embarking on EIM for the first time and government agencies may need to extend that initial time-to-value to about six to 12 months. "In the commercial world, you're trying to find one or two people who can say 'yes,'" he said. "In the federal world, there are, like, 1,000 people who can say 'no.'"

Forgetting the 80/20 rule
Organizations taking on an EIM program often do so under pretense that they need to "go after" every scrap of data in the company. But this is a mistake. According to Marco, about 80% of the value of

information comes from about 20% of the data housed within an organization, so companies should identify and focus their EIM efforts on the most important data first.

"Worry about managing the key data in the organization because that is going to give you the majority of the benefit from EIM efforts," he said.

Not ensuring ongoing adherence to maintenance

Taken as a whole, EIM should not be looked at as a project with an end in sight. Rather, Marco said, EIM is an ongoing initiative that requires ongoing maintenance -- a point that many organizations seem to forget. "A lot of times organizations focus so heavily on initially implementing their EIM effort that they forget [they] need to make sure that people adhere to the rules [they] create," he said. "Embed into your project development lifecycle as many of those EIM processes that you can." Marco added that it's a good idea to let data stewards -- the business users charged with enforcing EIM policies -- know that their bonuses, raises and other incentives will reflect how well they encouraged ongoing maintenance and adherence to EIM rules.

No active involvement by the business side

Business users understand the value of the information they create better than anyone, and therefore they must be closely involved in EIM efforts from the beginning. "I always know an EIM effort is heading for rocky waters when I [first visit a company and] meet a bunch of technologists," Marco said. "I need some people from the business, especially in… governance. That should be predominantly [business-driven]."

Neglecting metadata management

Metadata gives context to information. A retail establishment's product code for a can of Coca-Cola represents one piece of data. But information about the store where that can of soda came from and the town where that store resides are examples of context, or metadata. Without properly managed metadata, individual pieces of information can be largely useless.

"Metadata management and governance [efforts] need to begin before you begin your [large-scale] EIM work," Marco said.

Failing to change the status quo

Organizations need to clearly identify the problems that EIM can fix. There's no point in going on with an EIM effort if a company is still going to have a four-fold redundancy in data, Marco said. The projects must be designed to deliver real changes within a set amount of time.

"We're not doing EIM so that 5% of our data is in error," Marco said. "We're doing it to make things better. Our goal needs to be to give us a competitive differentiator from the marketplace."

Thinking that EIM is easy

Conference attendee Lisa Jane Bonamo, a data systems developer for Chicago-based insurance provider Health Care Service Corporation (HCSC), understands that enterprise information management efforts are far from easy. HCSC worked with Marco and Enterprise Warehousing Solutions to implement an EIM program that, among other things, is helping the organization manage its metadata environment and couple that environment with HCSC's enterprise logical data model and content management-related efforts. In doing so, Bonamo got to experience first-hand just how thorny an EIM project can be.

"It's difficult because there are so many complexities that you have to deal with to get that information into your tool correctly. You're not just dealing with one application and [you're] talking with many different people," Bonamo said. "You have the IT people who have their own way of doing things from a

technical view, and you're dealing with the governance people who are dealing with things from a personal point of view. You have to merge those things together and [it's] like a marriage. You have to make sure that everyone gets a turn at the table and everyone gets an opportunity to voice their opinions."