Prepared by:

**Poplin Working Group**

**MITA Reference Architecture**

Service Definition: Data Service Taxonomy

Initial Draft

Version 0.3

April 9, 2018

Record of Changes

| Version | Date | Author / Owner | Description of Change |
| --- | --- | --- | --- |
| 0.1 | 1/11/2018 | Dwayne Carter | Initial draft |
| 0.2 | 3/16/2018 | Dwayne Carter | Updated based on Poplin Group Feedback |
| 0.3 | 4/9/2018 | Dwayne Carter | Updated document with additional Poplin group feedback. |

Table of Contents

[Record of Changes 1](#_Toc511495687)

[Table of Contents 2](#_Toc511495688)

[1. Introduction 4](#_Toc511495689)

[1.1 Purpose 4](#_Toc511495690)

[1.2 Scope 4](#_Toc511495691)

[2. MITA Metamodel 4](#_Toc511495692)

[2.1 Purpose 4](#_Toc511495693)

[2.2 Model Diagram 5](#_Toc511495694)

[2.3 Components 5](#_Toc511495695)

[2.3.1 Business Architecture Components 5](#_Toc511495696)

[2.3.2 Information Architecture Components 6](#_Toc511495697)

[2.3.3 Technical Architecture Components 6](#_Toc511495698)

[3. MITA Metamodel: Data Service Extension 6](#_Toc511495699)

[3.1 Purpose 6](#_Toc511495700)

[3.2 Model Diagram 7](#_Toc511495701)

[3.3 Components 8](#_Toc511495702)

[4. Data Service Taxonomy 9](#_Toc511495703)

[4.1 Data Management Business Area Taxonomy 9](#_Toc511495704)

[4.1.1 Data Governance Management Category 10](#_Toc511495705)

[4.1.2 Data Architecture Management Category 10](#_Toc511495706)

[4.1.3 Data Modeling & Design Management Category 11](#_Toc511495707)

[4.1.4 Data Storage Management Category 11](#_Toc511495708)

[4.1.5 Data Security Management Category 11](#_Toc511495709)

[4.1.6 Data Integration & Interoperability Management Category 11](#_Toc511495710)

[4.1.7 Document & Content Management Category 12](#_Toc511495711)

[4.1.8 Reference Data Management Category 12](#_Toc511495712)

[4.1.9 Master Data Management Category 12](#_Toc511495713)

[4.1.10 Data Warehouse Management Category 13](#_Toc511495714)

[4.1.11 Business Intelligence Management Category 13](#_Toc511495715)

[4.1.12 Data Quality Management Category 13](#_Toc511495716)

[4.1.13 Metadata Management Category 14](#_Toc511495717)

[4.2 Data Processing Service Taxonomy 14](#_Toc511495718)

[4.2.1 Data Collection Service Classification 17](#_Toc511495719)

[4.2.2 Data Validation Service Classification 17](#_Toc511495720)

[4.2.3 Data Transformation Service Classification 18](#_Toc511495721)

[4.2.4 Data Storage Operations Service Classification 18](#_Toc511495722)

[4.2.5 Data Aggregation Service Classification 19](#_Toc511495723)

[4.2.6 Data Analysis Service Classification 19](#_Toc511495724)

[4.2.7 Data Presentation Service Classification 20](#_Toc511495725)

[4.2.8 Data Distribution Service Classification 20](#_Toc511495726)

[5. Data Service Taxonomy Use 21](#_Toc511495727)

[5.1 Data Management Taxonomy Use 21](#_Toc511495728)

[5.1.1 Data Management Business Services Identification 21](#_Toc511495729)

[5.1.2 Tools & Resources Standardization 21](#_Toc511495730)

[5.1.3 Data Management Strategy Organization 21](#_Toc511495731)

[5.2 Data Processing Taxonomy Use 22](#_Toc511495732)

[5.2.1 Data Transactions Development 22](#_Toc511495733)

[5.2.2 Reuse Opportunity Identification 22](#_Toc511495734)

[5.2.3 Target Schema Development 22](#_Toc511495735)

[5.2.4 Data Flow Development 23](#_Toc511495736)

[5.2.5 Standardize Data Processing 23](#_Toc511495737)

[5.2.6 Legacy Systems/Applications Functionality Identification 23](#_Toc511495738)

[5.2.7 Data Services Identification 23](#_Toc511495739)

[5.2.8 Data Service Capability Identification 24](#_Toc511495740)

[6. Works Cited 24](#_Toc511495741)

# Introduction

Organizations can’t escape the need to manage data. Organizations large and small must manage data to survive. One of the biggest differences between organizations that are successful and those that are not, is their ability to effectively manage and use data. One common misunderstanding is that the management of data is purely a function of technology. This often leads to business staff delegating all responsibility for the management of data to technical staff. Although technology often plays a major role in the management of data, it is not purely a function of technology. Data is a critical enterprise asset, that should be managed as such. Technology can help support the management of data but with or without technology, data must be managed.

## Purpose

The MITA 3.0 framework provides extensive guidance for State Medicaid Agencies to develop Business and Technical Services. Although the Information Architecture provides guidance on managing data, it lacks any discussion or guidance on the development of Data Services. The introduction of Data Services into the MITA framework will help provide additional opportunities for State Medicaid Agencies to collaborate and share both internally as well as with the vendor community. The purpose of this paper is to provide guidance for State Medicaid Agencies to develop Data Services that are aligned to the MITA 3.0 framework.

## Scope

Like Business and Technical Services, Data Services should be considered a part of the Technical Architecture. The process for developing a Data Service is very similar to the process for developing Business and Technical Services. The primary difference with Data Services is the functional taxonomy that helps with the initial planning and design. Business Services are driven by Business Processes and their associated Business Capabilities. Technical Services are driven by Technical Functions and their associated Technical Capabilities. The scope of this paper will focus on the following:

* + - Establish a standard definition for Data Service
    - Introduce Data Service into the MITA 3.0 Metamodel
    - Establish a functional taxonomy for classifying Data Services
    - Identify benefits and uses of the functional taxonomy
    - Identify next steps for expanding Data Services

# MITA Metamodel

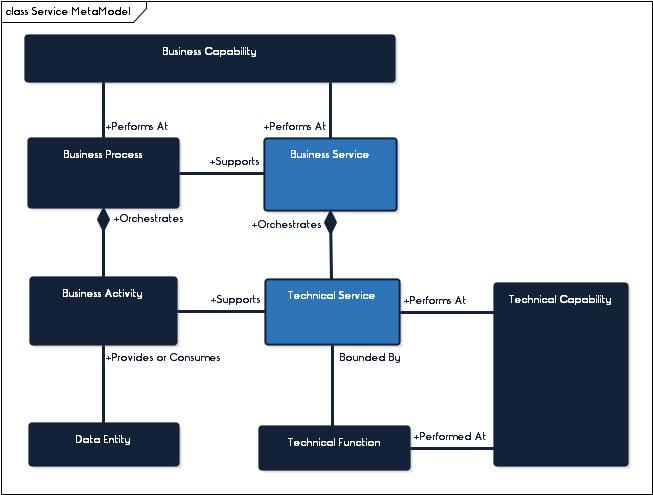
## Purpose

The MITA framework is an Enterprise Architecture framework designed to support State Medicaid Agencies. A critical component of developing any Enterprise Architecture is the development of an Architectural Metamodel. The Architectural Metamodel describes the critical components that are actively managed by an enterprise, as well as the relationships between those components. The Metamodel helps ensure that an enterprise is accurately developing models, designs and plans that is in the best interest of the entire Enterprise.

## Model Diagram

The following diagram identifies a high-level metamodel that describes the current relationships in the MITA Framework for Business and Technical Services.

Figure - MITA Metamodel



## Components

The following section will outline the components identified in ***Figure 1 - MITA Metamodel***. The components are organized by MITA Architecture which includes Business, Information and Technical Architectures.

### Business Architecture Components

* **Business Process –** The business process represents a process that is performed by a State Medicaid Agency to help meet its goals, objectives, and performance measures. Each business process is comprised of a set of smaller business activities that are performed to support the business process. The business process defines the orchestration of how and when these smaller business activities are performed.
* **Business Activity –** The business activity represents an individual step within a business process. Each activity performed, produces a meaningful result that supports a business process.
* **Business Capability –** A Business Capability is the competence of an individual, organization or system to perform a business process.
* **Business Service –** A Business Service is a software component that executes a business process or function at a specific capability level. It includes a defined interface for its invocation, performs a defined function that corresponds to the capability, and returns defined results.[[1]](#footnote-2) Business activities associated with processing data produce, manipulate and consume Data.

### Information Architecture Components

* **Data Entity –** A Data Entity represents an encapsulation of data that is recognized by the business as a thing. It is a collection of data that can represent any data set that an Enterprise manages. Data Entities are consumed by and/or produced by business activities. As an example. the business activity for “Receive Business Relationship Request” is a consumer for the “Business Relationship Request” Data Entity.

### Technical Architecture Components

* **Technical Functions –** Technical Functions are generic technical support activities. They are used to drive Technical Capabilities and define Technical Services.
* **Technical Service -** A Technical Service as a piece of software that executes a generic IT capability. It has a defined interface for its invocation, performs a defined function that corresponds to the capability, and returns defined results. [[2]](#footnote-3)
* **Technical Capabilities –** A Technical Capability is the competence of an individual, organization or system to perform a technical function.

# MITA Metamodel: Data Service Extension

## Purpose

Data Services represent a Data Transaction that has been automated by technology. It is a software component that executes a generic Data Processing capability. It has a defined interface for its invocation, performs a defined Data Processing Function that corresponds to the capability and returns defined results. The data service extension is intended to incorporate data services into the MITA framework. The purpose of the extension is to identify common data processing activities that can be reused across multiple business processes. These common data processing activities can be automated through data services and the applications and technology that support the data services.

## Model Diagram

The following diagram introduces the concept of Data Services into the MITA Metamodel.

Figure – MITA Data Services Extension

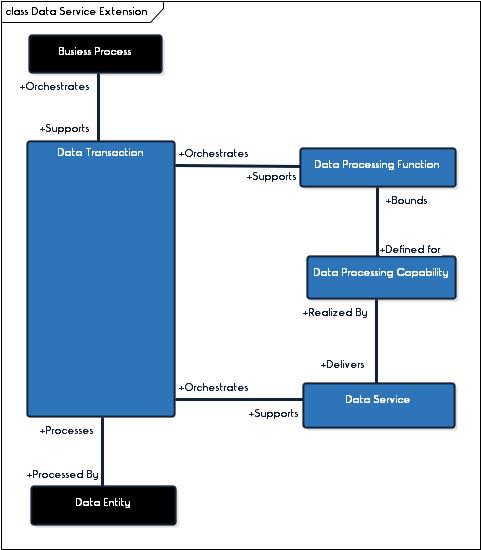
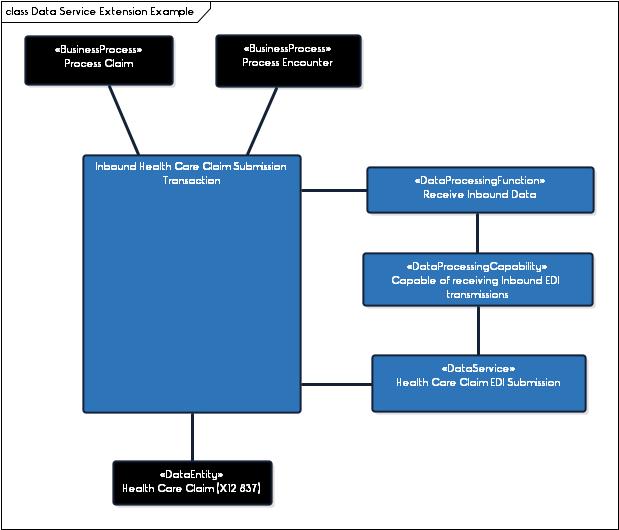


Figure - MITA Data Services Extension Example - Health Care Claim Submission



## Components

* + - **Data Processing Function –** A Data Processing Function represents generic activities or operations that are responsible for collecting and manipulating data, as well as converting that data into meaningful information that is used to support business goals and objectives. Data processing functions are generic and can be used to support a variety of different business processes.
    - **Data Transaction –** A Data Transaction represents a sub-process that orchestrates common data processing activities/operations. These data processing activities/ operations are specific to the transaction and realize data processing functions. A single data transaction can be used to support one or many different business processes. A single data transaction may also realize one or more data processing functions.
    - **Data Services -** A Data Service represents a Data Transaction that realizes a data processing capability. It is a software component that executes a generic Data Processing capability. Data Services have defined interfaces for their invocation and can perform a defined Data Processing Function that corresponds to the capability and returns defined results.
    - **Data Capability –** A Data Capability is the competence of an individual, organization or technology to perform a data processing function.

# Data Service Taxonomy

Data is the representations of facts about the world.[[3]](#footnote-4) Those facts are then interpreted by people and technology. The use of data enables an organization to fulfill its mission, to meet its goals and objectives. As technology has advanced, the amount of data available for an organization to use has substantially increased, making the appropriate management of data critical to organizations. Since data provides the connection from business to technology, it is important that the management meets the needs of both business and technology.

To illustrate these different views, the following sections introduce two taxonomies to support the management of data including a Data Management Business Area Taxonomy and a Data Processing Service Area Taxonomy. The Data Management Business Area Taxonomy is closely aligned to MITA Business Architecture and the Data Processing Service Area Taxonomy is closely aligned to the MITA Technical Architecture.

## Data Management Business Area Taxonomy

Data Management refers to the processes and activities associated with managing data as a valuable enterprise asset. Data Management represents the administrative side of managing data. It establishes strategies, policies, processes, performance metrics, etc. that support the management of data. The MITA 3.0 framework identifies a single process for “manage data” under the Operations Management Business Area. It also identifies the need to develop a Data Management Strategy under the Information Architecture section of the framework. To help drive the standardization of data management and increase interoperability, Data Management should be incorporated into the MITA business process hierarchy as a business area.

The MITA business process hierarchy is a structure that groups together business processes that have a common purpose and share data. The MITA business process hierarchy is organized into the following (3) tiers:

* + - Tier 1 is the business area grouping
    - Tier 2 is the business category grouping
    - Tier 3 is the business process

The Data Management Business Area is focused on establishing business processes that support the management of data. The new business area should be aligned to Data Management best practices. The following diagram outlines the Data Management Business Area and associated business categories.

Figure Data Management Business Area



### Data Governance Management Category

The Data Governance Management Category is focused on processes established to exercise authority, control, and shared decision-making (planning, monitoring, and enforcement over the management of data assets. (DAMA Intenational, 2017) Processes within this category include but are not limited to the following:

* + - Manage Data Management Strategy
    - Manage Data Management Roles & Responsibilities
    - Manage Data Management Standards
    - Manage Data Management Policies
    - Manage Data Management Processes
    - Manage Data Management Issues and Problems
    - Manage Business Glossary
    - Manage Data Asset Values

### Data Architecture Management Category

The Data Architecture Management Category is focused on processes established to identify the data needs of the enterprise (regardless of structure) and designing and maintaining blueprints to meet those needs. Using master blueprints to guide data integration, control data assets, and align data investments with business strategy. (DAMA Intenational, 2017) Processes within this category include, but are not limited to the following:

* + - Manage Data Architecture Viewpoints
    - Establish Data Architecture Views
    - Manage Data Architecture Roadmap
    - Inquire Data Architecture Information
    - Manage Data Architecture Assessments

### Data Modeling & Design Management Category

The Data Modeling & Design Management Category is focused on processes established to discover, analyze, and scope data requirements, and then representing and communicating these data requirements in a precise form called the data model. (DAMA Intenational, 2017) Processes within this category include, but is not limited to the following:

* + - Manage Enterprise Data Model
    - Establish Conceptual Data Model
    - Establish Logical Data Model
    - Manage Physical Data Model

### Data Storage Management Category

The Data Storage Management Category is focused on the processes associated with the design, implementation and support of stored data to maximize its value. (DAMA Intenational, 2017) Processes within this category include, but are not limited to the following:

* + - Manage Data Store Implementation
    - Manage Data Store Modification
    - Manage Data Store Destruction
    - Process Operational Data Queries
    - Manage Data Store Conversion
    - Manage Data Store Performance

### Data Security Management Category

The Data Security Management Category is focused on the processes associated with defining, planning, development and execution of security policies and procedures to provide proper authentication, authorization, access, and auditing of data and information assets. (DAMA Intenational, 2017) Processes within this category include, but are not limited to:

* + - Manage Data Security & Privacy Rules
    - Manage Data Security & Privacy Classifications
    - Manage Data Releases
    - Manage Data De-Identification
    - Manage Data Security and Privacy Controls & Procedures
    - Manage Data Security and Privacy Assessments

### Data Integration & Interoperability Management Category

The Data Integration & Interoperability Management Category is focused on the processes associated with moving and consolidating data within and between applications and organizations. (DAMA Intenational, 2017) Processes within this category include, but are not limited to the following:

* + - Manage External Data Sources
    - Establish Data Flow Models
    - Manage Data Harmonization
    - Process Data Acquisition Requests
    - Determine Best Data Store Source
    - Determine Best Data Store Target
    - Manage Inbound Data Transactions
    - Manage Outbound Data Transactions
    - Manage Inbound Paper Document Submissions
    - Manage Outbound Paper Document Distributions
    - Manage Data Subscriptions and Publications
    - Manage Data Publishing
    - Manage Data Extraction, Transformation and Loading (ETL)
    - Manage Data Integration Information Inquiries

### Document & Content Management Category

The Document and Content Management Category is focused on the processes associated with the planning, implementation, and control of data and information found in any form or medium. (DAMA Intenational, 2017) Processes within this category include, but are not limited to the following:

* + - Maintain Document and Record Data Store
    - Manage Document and Record Generation
    - Manage Document and Record Conversion
    - Manage Document and Record Destruction
    - Manage Content Creation
    - Manage Content Modification
    - Manage Content Destruction
    - Manage Document and Content Information Inquiries

### Reference Data Management Category

The Reference Data Management Category is focused on the processes associated with managing reference information to meet organizational goals, reduce risks associated with data redundancy, ensure high quality and reduce the costs of data integration. (DAMA Intenational, 2017) Processes within this category include, but are not limited to the following:

* + - Maintain Reference Data Store
    - Manage Reference Data Sets
    - Manage Reference Data Values
    - Manage Reference Data Codes
    - Manage Reference Data Modifications
    - Manage Reference Data Removals
    - Maintain External Reference Data Standards
    - Manage Reference Data Information Inquiries

### Master Data Management Category

The Master Data Management Category is focused on the processes associated with linking critical data to a central master file that is used as reference information to meet organizational goals, reduce risks associated with data redundancy, ensure high quality and reduce the costs of data integration. (DAMA Intenational, 2017) Processes associated with this category include, but are not limited to the following:

* + - Maintain Master Data Store
    - Manage Master Data Rules
    - Manage Master Data Sources
    - Determine Master Data Record Matches
    - Manage Master Data Record Creation
    - Manage Master Data Record Modifications
    - Manage Master Data Record Deletion
    - Manage Master Data Inquiries

### Data Warehouse Management Category

The Data Warehouse Management Category is focused on the processes associated with the planning, implementation, and controls that provide decision support data to the organization. (DAMA Intenational, 2017) Processes associated with this category include, but are not limited to the following:

* + - Maintain Data Warehouse (Analytical Data Store)
    - Maintain Data Mart (Analytical Data Store)
    - Manage Analytical Data Sources
    - Process Analytical Data Queries
    - Manage Analytical Data Dimensions
    - Manage Analytical Data Facts & Measures

### Business Intelligence Management Category

The Business Intelligence Management Category is focused on the processes associated with the collection and analysis of many different types of data to support reporting, query, analysis and to find answers and insights for questions that are not known at the start of analysis. (DAMA Intenational, 2017) Processes within this category include, but are not limited to the following:

* + - Manage Business Intelligence Rules
    - Manage Data Gathering and Preparation
    - Manage Hypothesis Testing
    - Process Descriptive Analytics Requests
    - Process Predictive Analytics Requests
    - Process Prescriptive Analytics Requests
    - Process Diagnosis Analytics Requests
    - Manage Reporting
    - Manage Data Visualization
    - Manage Dashboards

### Data Quality Management Category

The Data Quality Management Category is focused on the processes associated with the planning, implementation, and control of activities that apply quality management techniques to data, in order to assure it is fit for consumption and meets the needs of data consumers. (DAMA Intenational, 2017) Processes within this category include, but are not limited to the following:

* + - Manage Data Quality Rules
    - Manage Data Quality Dimensions
    - Process Data Element Quality Assessment Requests
    - Process Data Record Quality Assessment Requests
    - Process Data Set Quality Assessment Requests
    - Manage Data Quality Issues
    - Manage Data Cleansing
    - Manage Data Enhancements
    - Manage Data Quality Information Inquiries

### Metadata Management Category

The Metadata Management Category is focused on the processes associated with planning, implementation, and control activities to enable access to high quality and integrated metadata. (DAMA Intenational, 2017) Processes within this category include, but are not limited to the following:

* + - Maintain Metadata Store
    - Manage Metadata Sources
    - Establish Metamodel
    - Manage Metadata Creation
    - Manage Metadata Modification
    - Manage Metadata Removal
    - Manage Metadata Information Inquiries

## Data Processing Service Taxonomy

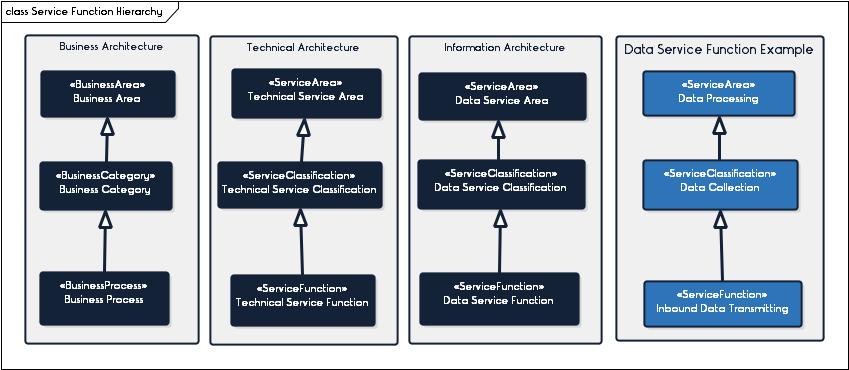
Data Processing is facilitated and governed by Data Management. It refers to the collection and manipulation of data to produce meaningful information that is then delivered and distributed for business use. All business processes require some form of data processing. Data Processing activities are focused on the following:

* + - Bringing data into an organization
    - Ensuring that data meets the needs of the organization
    - Changing the data to meet the needs of an organization
    - Storing and retaining the data for an organization to use
    - Producing information to help support better understanding and decision making
    - Reporting the data to organization stakeholders

The processing of data to support business processes can be considered a Data Service. Many Data Services are reusable across business processes and can support the MITA goals of increasing interoperability. Like Business and Technical Services, Data Services could be defined at various levels of granularity. Each service is driven by taxonomies that have three (3) levels of granularity. The Business Service has Business Service Areas, Categories and Processes. The Technical Service has Technical Service Areas, Categories and functions. To ensure alignment, Data services will also be defined at three (3) levels of granularity which include the following:

* + - **Tier 1: Data Service Areas** represent the highest-level of granularity used to describe a Data Service Function.
    - **Tier 2: Data Service Category** represents an additional classification or grouping of Data Service Functions.
    - **Tier 3: Data Service Function** represents the lowest-level for describing a Data Service Function.

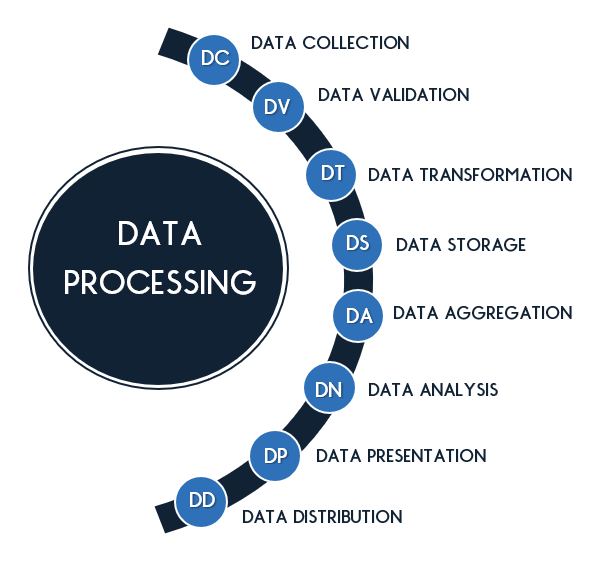
Figure Service Function Hierarchy



**Taxonomy Details**

The Data Processing Service Area is focused on identifying the generic data processing functions that are performed by individuals or technology. Each data processing function can be aligned to a single or group of business activities that are performed to support a business process. The Data Processing Service Area has eight (8) functional categories that are aligned to the Data Processing Lifecyle.

Figure – Data Processing Cycles



The following sections will provide details for each of the Data Processing Service Areas. Each section will include a table that outlines information about the service functions found within each Service Area including the following:

* + - **ID –** Represents a unique identification number assigned to the Service Function
    - **Service Function –** Represents the name of the Service Function.
    - **Service Classification –** Represents the Data Processing Service Area that classifies the Service Function.
    - **Service Function Description –** Represents a definition for the Service Function
    - **Common Aliases –** Represents common names that are used to describe the Service Function.

### Data Collection Service Classification

The Data Collection Service Classification (DC) is focused on bringing data into an organization. All functions in this category play a role in receiving data and successfully bringing data into an organization for use.

Table – Data Collection Service Functions

| ID | Service Function | Service Function Description | Common Aliases |
| --- | --- | --- | --- |
| DC-01 | Capture data | Function for capturing machine readable text. The capturing can happen manually through data entry or converting a digital image into machine readable text. | Data Scraping |
| DC-02 | Receive structured data | Function for receiving inbound structured data that has been sent from data trading partners. The structured data refers to data that is highly structured in a fixed field within a record or file and can be converted into a data model. Structured data is most commonly associated with storage in a relational database. |  |
| DC-03 | Receive unstructured data | Function for receiving inbound unstructured data that has been sent from data trading partners. The unstructured data refers to data that does not have a common structure. Unstructured data is most commonly associated with photos, graphic images, PDF files, paper documents, emails, etc. |  |
| DC-04 | Receive semi-structured data | Function for receiving semi-structured data that has been sent from data trading partners. The semi-structured data refers to data that has a structure but can’t be converted into a data model. Semi-structured data relies on tags and other types of markers used to identify elements within the data. Semi-structured data is most commonly associated with XML, JSON, etc. |  |

### Data Validation Service Classification

The Data Validation Service Classification (DV) is focused on measuring and testing the quality of data to determine if the data conforms to the applicable conditions or specified requirements. The service classification is focused on validating the quality of data and not resolving any data quality issues.

Table – Data Validation Service Functions

| ID | Service Function | Service Function Description |
| --- | --- | --- |
| DV-01 | Validate Data Accuracy | Function for validating that the data correctly represents ‘real life’ |
| DV-02 | Validate Data Completeness | Function for validating that all required data is present. |
| DV-03 | Validate Data Consistency | Function for validating that data is represented consistently within, between and across data sets. |
| DV-04 | Validate Data Integrity | Function for validating that data is aligned to the expected structure |
| DV-05 | Validate Data Reasonability | Function for validating that data and data patterns meet expectation. |
| DV-06 | Validate Data Timeliness | Function for validating that the timing associated with the data meets expectations and needs |
| DV-07 | Validate Data Uniqueness | Function for validating that no entry in the data exists more than once and there are no duplicates |
| DV-08 | Validate Data Validity | Function for validating that the data is consistent with expected reference data values and/or codes |

### Data Transformation Service Classification

The Data Transforming Service Classification (DT) is focused on changing data. This includes semantic, structural, or other changes to help improve the use and/or quality of the data.

Table – Data Transformation Service Functions

| ID | Service Function | Service Function Description |
| --- | --- | --- |
| DT-01 | Augment Data | Function for adding to the data such as demographic data, time/date data, etc. to enhance the data. |
| DT-02 | Replace Data | Function for removing data and replacing it with other data. |
| DT-03 | De-duplicate Data | Function for removing duplicative data that is repeated and not needed within a data set and/or record. |
| DT-04 | Remove Data | Function for removing data that is not needed within a data set and/or record. |
| DT-05 | De-Identify Data | Function for removing any identifiers that can be traced to an individual person, organization, etc. |
| DT-06 | Classify (Code/Reference) Data | Function for adding reference data to help classify and standardize data. |
| DT-07 | Restructure data | Function for changing the structure of data to match the target. |
| DT-08 | Image Document | Function for scanning and converting paper documents into Digital Images. |
| DT-09 | Map Data | Function for relating one data element to another data element to signify that the two data elements are related. Function provides connection between data elements in a source schema to data elements in a target schema. |
| DT-10 | Generate Document | Function for transforming data (including content) into a document. |

### Data Storage Operations Service Classification

The Data Storage Operations Service Classification is focused on data operations within a data store. This includes operations on a variety of different types of data including transaction data, analytical data, metadata, master data, etc.

Table – Data Storage Service Functions

| ID | Service Function | Service Function Description |
| --- | --- | --- |
| DS-01 | Create Data | Function for creating data in a data store. This includes data objects, records, elements, etc. |
| DS-02 | Delete Data | Function for deleting and removing data from a data store. |
| DS-03 | Extract Data | Function for gathering and removing data from a data store. |
| DS-04 | Insert Data | Function for inserting data into an existing data store. |
| DS-05 | Load Data | Function for gathering and inserting batches of data into a data store. |
| DS-06 | Retrieve Data | Function for searching, finding and reading data in an existing data store. |
| DS-07 | Update Data | Function for updating existing data in a data store. |
| DS-08 | Link Data | Function for linking different records that have an association with each other. |
| DS-09 | Merge Data | Function for merging different data records that are similar into one comprehensive record. |
| DS-10 | Destroy Document | Function for destroying paper documents stored in a physical document storage container. *(i.e. File Cabinet)* |
| DS-11 | File Document | Function for inserting paper documents into a physical document storage container. *(i.e. File Cabinet)* |
| DS-12 | Extract Document | Function for searching and extracting paper documents from a physical document cabinet container. *(i.e. File Cabinet)* |

### Data Aggregation Service Classification

The Data Aggregation Service Classification (DA) is focused on bringing data together to support data analysis. This includes the development of different views of the data based on user needs.

Table – Data Aggregation Service Functions

| ID | Service Function | Service Function Description |
| --- | --- | --- |
| DA-01 | Dice Data | Function for developing analytical views of data that crosses multiple analytical data dimensions. |
| DA-02 | Drill-Down Data | Function for developing analytical views of data that move from general views of the data to more granular layers of the data. |
| DA-03 | Group Data | Function for developing analytical views of data that grouping together data and create a summary view of the group. |
| DA-04 | Pivot Data | Function for developing analytical views of data that groups data by different analytical data dimensions and provides defined calculations by dimension including things such as sum, averages, frequency, etc. |
| DA-05 | Roll-Up Data | Function for developing analytical views of data that move from granular levels of data to more generic summary views. |
| DA-06 | Slice Data | Function for developing analytical views of data that spans an analytical data dimension. |

### Data Analysis Service Classification

The Data Analysis Service Classification(DN) is focused on turning raw data into meaningful information that helps improve decision making.

Table – Data Analysis Service Functions

| ID | Service Function | Service Function Description |
| --- | --- | --- |
| DN-01 | Profile Data | Function for gathering descriptive statistics about data that can help detect anomalies about the data. |
| DN-02 | Associate Data | Function for identifying data within a data set that frequently appear together and appear to have an association. |
| DN-03 | Classification Data | Function for identifying patterns by assigning items to classification groupings to predict how other similar items will be classified or grouped. |
| DN-04 | Cluster Data | Function for finding natural groups of data within a data set. |
| DN-05 | Data Estimation | Function for using sample data and statistical models to identify estimates about a larger population. |
| DN-06 | Data Optimization | Function for finding the best possible value to a given problem, given a set of defined constraints. Function uses a variety of different techniques to find the best available value. |
| DN-07 | Data Regression | Function for estimating the relationships among variables. |
| DN-08 | Sequence Data | Function for finding sequence patterns in each data set. |
| DN-09 | Data Simulation | Function for using statistical models to simulate the real world and make predictions. |
| DN-10 | Statistical Modeling | Function for developing models from a data sample that can be generalized to a larger population. There are several different types of Statistical Modeling including predictive models, regression models, etc. |
| DN-11 | Time-Series Forecasting | Function for forecasting future events based on assessing past events that include a timing component. |

### Data Presentation Service Classification

This Data Presentation Service Classification (DP) is focused on converting data into a format that can be presented to stakeholders.

Table – Data Presentation Service Functions

| ID | Service Function | Service Function Description |
| --- | --- | --- |
| DP-01 | Dashboard Data | Function for inputting data into a Dashboard. |
| DP-02 | Report Data | Function for generating a report from data. |
| DP-03 | Visualize Data | Function for generating a graphic representation of data. |

### Data Distribution Service Classification

The Data Distribution Service Classification (DD) is focused on distributing data out of an organization. All functions in this category play a role in sending data and successfully delivering it out of an organization for use.

Table – Data Distribution Service Functions

| ID | Service Function | Service Function Description |
| --- | --- | --- |
| DD-01 | Send structured data | Function for sending outbound structured data to data trading partners. The structured data refers to data that is highly structured in a fixed field within a record or file and can be converted into a data model. Structured data is most commonly associated with storage in a relational database. |
| DD-02 | Send unstructured data | Function for sending outbound unstructured data to data trading partners. The unstructured data refers to data that does not have a common structure. Unstructured data is most commonly associated with photos, graphic images, PDF files, paper documents, emails, etc. |
| DD-03 | Send semi-structured data | Function for sending outbound semi-structured data to data trading partners. The semi-structured data refers to data that has a structure but can’t be converted into a data model. Semi-structured data relies on tags and other types of markers used to identify elements within the data. Semi-structured data is most commonly associated with XML, JSON, etc. |
| DD-04 | Publish Data | Function for distributing data to a group of subscribers that have subscribed to receive a data publication. |

# Data Service Taxonomy Use

## Data Management Taxonomy Use

The Data Management Business Area Taxonomy can be used to help SMA’s meet their goals. The following section identifies how the taxonomy can be used to support the SMA.

### Data Management Business Services Identification

The Data Management Business Area Taxonomy can be used to drive the development of Data Management Business Services. Like other processes within the MITA Framework, each data management business process will define a different business service. These data management business services can be reused and shared within and across organizations.

### Tools & Resources Standardization

The Data Management Taxonomy can be used to identify tools and resources that support each process/category and or area. Once identified, these can be leveraged to help support other areas within the organization that are performing similar processes but have no tool or resources.

### Data Management Strategy Organization

The Data Management Taxonomy can be used to drive the SMA Data Management Strategy. Each Data Management Business Category represents a significant function of data management. As such, each category should have a fully developed strategy that identifies how the SMA will manage that function. Together all functional strategies define the SMA Data Management Strategy. This helps the SMA consider and develop an approach for managing all applicable data management functions and drive standardization across the enterprise.

## Data Processing Taxonomy Use

The Data Processing Service Area Taxonomy can be used to help SMA’s meet their goals. The following section identifies how the taxonomy can be used to support the SMA.

### Data Transactions Development

The Data Service Taxonomy can be used within the Business Architecture to support Business Process Modeling. Each business process identified by MITA includes a set of steps or activities that are performed to support that business process. Within each process, there are a number of these activities that can be grouped as a single transaction. Business Process Modeling Notation (BPMN) identifies a transaction as a specialized sub-process that requires all activities within the sub-process to be complete before the transaction is considered complete. The Data Service Taxonomy can be used to help define the Data Transactions within a business process. Each activity in a business process that is associated with processing data (i.e. sending, receiving, retrieving, etc.) can be grouped into a single Data Transaction using the Data Service Taxonomy. The following table identifies a list of possible Data Transactions using the Data Service Taxonomy.

Table – Data Transaction Table

| Function | Business Process | Business Activity | Transaction |
| --- | --- | --- | --- |
| Receive Unstructured Data | Establish Business Relationship | Receive request for business relationship from partner. | Business Relationship Request Submission |
| Receive Semi-structured data | Process Claim | Receive claim submission or claim adjustment information. | Health Care Claim Submission |
| Receive Semi-structured data | Process Encounter | Receive claim submission or claim adjustment information. | Health Care Claim Submission |
| Receive Unstructured Data | Establish Business Relationship | Receive request for business relationship from partner. | Business Relationship Request Paper Document Submission |

### Reuse Opportunity Identification

The use of the Data Service Taxonomy to define Data Transactions will provide a mechanism for State Medicaid Agencies to easily identify internal reuse opportunities. There are many Data Transactions, such as “Provider NPI Validation,” that are used by a variety of different business processes, departments, and units. Often, this is being handled differently by each one. The modeling of the Data Transaction will provide a tool for identifying collaboration opportunities. It also provides a tool for identifying plans to increase the overall enterprise maturity for managing data.

### Target Schema Development

Each Data Transaction identified will also be associated with a “Transaction Set” Data Entity. This can then be used to identify if there is an existing structure or transaction standard. It also provides a common vocabulary to work with internal, intrastate, interstate and regional partners to establish standards. The identification and standardization of these Transaction Data Set Schemas can be leveraged to increase interoperability, decrease time for data exchange mapping and support capability maturity.

Table – Data Transaction Set Table

|  |  |  |  |
| --- | --- | --- | --- |
| Function | Transaction | Transaction Set | Structure Standard |
| Receive Unstructured Data | Business Relationship Request Submission | Business Relationship Request | N/A |
| Receive Semi-structured data | Health Care Claim Submission | Health Care Claim | ASC X12 837 Health Care Claim |
| Receive Semi-structured data | Health Care Claim Submission | Health Care Claim | ASC X12 837 Health Care Claim |
| Receive Unstructured Data | Business Relationship Request Paper Document Submission | Business Relationship Request | N/A |

### Data Flow Development

Once a Data Transaction has been identified, a Data Flow that supports that transaction can be developed. The Data Flow helps identify data sources, targets, transmission and needs. It provides a data view for each business process that can be used to support capability maturity.

Table – Data Transaction Source/Target Table

|  |  |  |  |
| --- | --- | --- | --- |
| Function | Transaction | Source | Target |
| Receive Unstructured Data | Business Relationship Request Submission | Business Relationship Applicant | Inbound |
| Receive Semi-structured data | Health Care Claim Submission | Provider | Inbound |
| Receive Semi-structured data | Health Care Claim Submission | Provider | Inbound |
| Receive Unstructured Data | Business Relationship Request Paper Document Submission | Any Stakeholder | Inbound |

### Standardize Data Processing

The Data Service Taxonomy can be used standardize the functions that dictate how each business process performs these functions. It also allows for the ability to easily define and perform audits to ensure data is being processed as expected across the enterprise.

### Legacy Systems/Applications Functionality Identification

The Data Service Classification can be used by the Technical Architecture as reference data for classifying data processing functionality provided by current systems and applications. It will allow the Technical Architecture team to have a common vocabulary when discussing data processing technology.

### Data Services Identification

The Data Service Classification can be used by the Technical Architecture as reference data for classifying future Data Services, as well as identifying enterprise-wide requirements for those services. It can also significantly decrease the time for design if the business and information architecture teams have also used the Data Service Classification to define Data Transactions and Transaction Set Schemas.

### Data Service Capability Identification

The Data Service Taxonomy can be used to develop Information Capabilities that are closely aligned with the processing of data. It will alleviate confusion with the current capability model which does not easily connect to specific business processes such as MITA Business and Technical capabilities. It also allows for more detailed Data-related capabilities.

# Works Cited

CMS - Center for Medicare and Medicaid Services. (2012). MITA 3.0 - Part I Business Architecture. *MITA 3.0 Framework*.

CMS - Center for Medicare and Medicaid Services. (2012). MITA 3.0 - Part II Information Architecture. *MITA 3.0 Framework*.

CMS Center for Medicare and Medicaid Services. (2012). MITA 3.0 - Part III Technical Architecture. *MITA 3.0 Framework*.

DAMA Intenational. (2017). *DAMA - DMBOK Data Management Body of Knowledge 2nd Edition.* Basking Ridge: Technics Publications.

1. Data Management Business Process Template Example

The following table identifies an example Data Management Process described using a Business Process Template.

| DM-XX – Manage Inbound Information Submission | | | | |
| --- | --- | --- | --- | --- |
| Item | Details | | | |
| Description | The ***Receive Inbound Data Submission*** is a Data Management process for receiving inbound electronic information. Stakeholder prepare the information using a defined structure and format and submits the information to the State Medicaid Agency (SMA). The SMA receives the information and processes it for business use. This business process supports all State Medicaid Agencies that require the electronic submission of information.  **Note:** The specific process steps will vary by type of information. Each information type will include business rules that will drive the process. | | | |
| Trigger Event | Environment-based Trigger Events to include but not be limited to:   * Stakeholder prepares information and sends electronically to SMA | | | |
| Result | * SMA validates, transforms and stores information for business use * SMA sends acknowledgement of receipt to sender | | | |
| Business Process Steps | **#** | **Process Step** | | **Data Processing Service Function** |
| 1 | **Trigger:** Stakeholder sends information. | | *N/A* |
| 2 | Receives Information. | | Receive Inbound Electronic Data Transmission |
| 3 | Log Information | | Data Creation |
| 4 | Send alert to trigger ***Manage Data Quality Validation*** process to validate information.[Subprocess: ***Manage Outbound Information Distribution***] | | *[Subprocess]* |
| 5 | Send acknowledgement of information receipt notification [Subprocess: ***Manage Outbound Information Distribution***] | | *[Subprocess]* |
| 6 | Send request to trigger ***Determine Master Data Match*** process. [Subprocess: ***Manage Outbound Information Distribution***] | | *[Subprocess]* |
| 7 | If no Master Data Match, send request to trigger ***Determine Best Data Storage Target*** process. [Subprocess: ***Manage Outbound Information Distribution***] | | *[Subprocess]* |
| 8 | Transform data to match Target Schema. | | Data Restructuring |
| 9 | If Operational Information, send alert to trigger ***Manage Operational Data Transactions*** process to store operational data.[Subprocess: ***Manage Outbound Information Distribution***] | | *[Subprocess]* |
| 10 | If Analytical Information, send alert to trigger ***Manage Analytical Data Transactions*** process to store analytical data.[Subprocess: ***Manage Outbound Information Distribution***] | | *[Subprocess]* |
| 11 | Determine the authorized subscribers that need information. | | *N/A* |
| 12 | **End:** Send New “Information Available “alert to authorized subscribers. [Subprocess: ***Manage Outbound Information Distribution***] | | *[Subprocess]* |
| Primary Messages | **Primary Messages**   * Information Submission * Information Submission Successfully Notification * Information Submission Failed Notification | | | |
| **Secondary Messages** | * Data Validation Request * Data Validation Response * Master Data Match Request * Master Data Match Response * Target Data Storage Location Request | | * Target Data Storage Location Response * Operational Data Record Creation Request * Operational Data Record Creation Response * Analytical Data Record Creation Request * Analytical Data Record Creation Response | |
| Data Entities | * Information (Type based on process) * Information Submission Data (Transactional) * Information Transformation Data (Transactional) | | | |
| Successors | * Manage Data Quality Validation * Manage Outbound Information Distribution * Determine Master Data Match | | * Determine Best Data Storage Target * Manage Operational Data Transactions * Manage Analytical Data Transactions | |
| Failures | Information does not pass validation and sender receives a submission failed notification. | | | |
| Performance Measures | * Time to receive information and generate a response. \_\_\_\_\_Hours * Time to receive information and alert authorized subscribers of receipt. \_\_\_\_\_\_Hours * Time to transform information from source to target schema. \_\_\_\_\_\_Hours * Standard Response Notification Usage \_\_\_\_\_% of time | | | |

1. Data Management Business Capability Example

The following table provides an example Business Capability Matrix for the Data Management Process “Receive Inbound Data Submissions”.

| Capability | Level 1 | Level 2 | Level 3 | Level 4 | | Level 5 |
| --- | --- | --- | --- | --- | --- | --- |
| Receive Inbound Data Submission | | | | | | |
| Business Capability Quality: Timeliness of Process | | | | | | |
| How timely is the end-to-end process?  *SMA is capable of processing electronic Inbound Data Submissions in a timely manner* | ***More than 5%*** of Internal Inbound Data Submissions can be received electronically. | 100% of Internal Enterprise Inbound Data Submissions can be received electronically AND processed in ***10 minutes*** or less. | 100% of Internal Enterprise and 100% of Intrastate Inbound Data Submissions can be received electronically AND processed in ***5 minutes*** or less. | 100% of Internal Enterprise, Intrastate, and Regional Inbound Data Submissions can be received electronically AND processed in **2 *minutes*** or less. | 100% of Internal Enterprise, Intrastate, Regional, and National Inbound Data Submissions can be received electronically AND processed in **1 *minutes*** or less. | |
| Business Capability Quality: Data Access and Accuracy | | | | | | |
| How accurate is the information in the process?  *SMA can ensure that all data received during submission meets data accuracy expectations* | Submitted Data from internal enterprise inbound data submissions has an overall Data Accuracy rating of ***at least 60%*** | Submitted Data from internal enterprise inbound data submissions has an overall Data Accuracy rating of ***85% or more.*** | Submitted Data from internal enterprise and intrastate inbound data submissions has an overall Data Accuracy rating of **90*% or more.*** | Submitted Data from internal enterprise, intrastate, and regional inbound data submissions has an overall Data Accuracy rating of **95*% or more.*** | Submitted Data from internal enterprise, intrastate, regional, and national inbound data submissions has an overall Data Accuracy rating of **100*% or more.*** | |

1. Data Service Technical Capability Example

The following provides an example technical capability matrix for the Data Service Function “Receive Inbound Data Transmission”.

| Capability | Level 1 | Level 2 | Level 3 | Level 4 | Level 5 |
| --- | --- | --- | --- | --- | --- |
| Data Collection | | | | | |
| Receive Semi-Structured Data | | | | | |
| SMA can receive electronic data in real-time | Less than 90% of Internal Enterprise transactions that “Receive Inbound Electronic Data Transmission” occur in real-time | 90% of Internal Enterprise transactions that “Receive Inbound Electronic Data Transmission” occur in real-time | 90% of Intrastate transactions that “Receive Inbound Electronic Data Transmission” occur in real-time. | 90% of Regional transactions that “Receive Inbound Electronic Data Transmission” occur in real-time. | 90% of National transactions that “Receive Inbound Electronic Data Transmission” occur in real-time. |
| SMA can receive electronic batch data | No transactions associated with “Receive Inbound Electronic Data Transmission” uses batch processing. | At least (1) Internal Enterprise transaction associate with “Receive Inbound Electronic Data Transmission” uses batch processing. | At least (1) Intrastate transaction associate with “Receive Inbound Electronic Data Transmission” uses batch processing. | At least (1) Regional transaction associate with “Receive Inbound Electronic Data Transmission” uses batch processing. | At least (1) National transaction associate with “Receive Inbound Electronic Data Transmission” uses batch processing. |
| SMA can standardize the data syntax Inbound Electronic Data Transmissions | Less than 95% of Internal Enterprise transactions that “Receive Inbound Electronic Data Transmission” use an Enterprise Structure Data Standard. | At least 95% of Internal Enterprise transactions that “Receive Inbound Electronic Data Transmission” use an Enterprise Structure Data Standard. | At least 95% of Intrastate transactions that “Receive Inbound Electronic Data Transmission” use an intrastate Structure Data Standard. | At least 95% of regional transactions that “Receive Inbound Electronic Data Transmission” use a regional Structure Data Standard. | At least 95% of national transactions that “Receive Inbound Electronic Data Transmission” use a national Structure Data Standard. |
| SMA can standardize the data semantics of Inbound Electronic Data Transmissions | Less than 95% of Internal Enterprise transactions that “Receive Inbound Electronic Data Transmission” have been aligned to the Enterprise Vocabulary Data Standards, | At least 95% of Internal Enterprise transactions that “Receive Inbound Electronic Data Transmission” have been aligned to the Enterprise Vocabulary Data Standards, | At least 95% of Intrastate transactions that “Receive Inbound Electronic Data Transmission” have been aligned to Intrastate Vocabulary Data Standards, | At least 95% of Regional transactions that “Receive Inbound Electronic Data Transmission” have been aligned to Regional Vocabulary Data Standards, | At least 95% of National transactions that “Receive Inbound Electronic Data Transmission” have been aligned to National Vocabulary Data Standards, |

1. Chapter II – Business Services [↑](#footnote-ref-2)
2. Chapter II – Technical Services [↑](#footnote-ref-3)
3. DAMA [↑](#footnote-ref-4)