New Mexico Data Services Module

Data Models

Date of Issue: October 6, 2020
Deliverable: DS11
Revision Status: Approved with DDI Release 7
Revision: v7.0

A collaboration of

HHS 2020 Person-Centered Health and Human Services

IBM Watson Health

Table of Contents

List of Tables

Document Information

Document Identification

Document Identification	
Deliverable Title	Data Models
Contract Reference	Exhibit A, Section IV., C. Deliverable 11
IBM Owner	Todd Smith
HSD Owner/Approver	Mike Norero
HSD Reviewers	Manny Martinez, Jennifer Galvez, Joe Jones, Jan-Marie Guzman, Kristie Abraham, Mark Pitcock, Mario Gonzalez

Revision History

Date	Version	Change Description
07/01/2019	1.0	Submitted to HSD for review
07/31/2019	1.1	With IBM responses to HSD comments on v1.0
08/30/2019	1.2	With IBM responses to HSD comments on v1.1
09/30/2019	1.3	With IBM responses to HSD comments on v1.2
10/15/2019	2.0	Final approved deliverable
11/15/2019	3.0	Updated for DDI-4. Physical and logical data models for Provider, Client, MCO, TPL, Finance, Drug Rebate added. EEM data models for Finance and TPL added
12/17/2019	3.1	Conceptual models for Provider, Client, MCO, TPL, Finance, and Drug Rebate added
01/21/2020	3.2	Replaced TPL STG file; revised logical data models
02/03/2020	4.0	Final approved by HSD Jan 31, 2020, as part of DDI Release 4
02/07/2020	4.1	Updated as part of DDI Release 5. Added Client data models in Section 7.3.
03/11/2020	4.2	Responses to HSD comments for resubmittal with DDI Release 5. as well as updates related to findings during internal review
03/30/2020	5.0	Final, accepted version with DDI Release 5
05/06/2020	5.1	Removed links to individual files from the Diagrams table in Sections 6.3, 7.3, and 8.2 and validated links to the SharePoint folders that hold the files. Added Client and Provider EEM data models as part of DDI Release 6.

Date	Version	Change Description
		Added new and updated MCO data models for multiple schemata in Sections 6.3, 7.3, and 8.2. Updated content in Sections 7.1 and 8.1; added new Flexible Analytics data models for multiple schemata in Sections 7.3 and 8.2 only, as Flexible Analytics has does not have a conceptual data model
06/05/2020	5.2	Updated Sections 7.1, 7.3, 8.1, and 8.2 to address HSD request for additional context around the addition of Flexible Analytics data models
06/24/2020	6.0	Final, accepted version with DDI Release 6
07/29/2020	6.1	Updates for DDI Release 7: Data Models in Section 8.2: Diagrams
08/26/2020	6.2	Updates to address HSD v6.1 comments and collaboration session
10/06/2020	7.0	Approved by HSD

Related Documents

The following are Data Services module deliverables related to the Data Model deliverable.

Deliverable Name
Metadata Catalog
Master Test Plan
Resource Management Plan and Staffing Addendum
Risk Management Log Addendum
Security Design Plan
System Design Document

1 INTRODUCTION

1.1 Overview

The Data Services (DS) Data Models deliverable documents the data modeling approach and data models of the DS module for the New Mexico Human Services Department (HSD) Medicaid Management Information System Replacement (MMISR) project.

The Data Models will be developed based on the following:

1. It will comply with the System Integrator (SI)-owned data models deliverables and the standards established therein
2. It will comply with the Centers for Medicare & Medicaid Services (CMS) eXpedited Lifecycle (XLC) Template for the Logical Data Model (LDM)

1.2 Project Scope

The New Mexico HSD has adopted the Health and Human Services (HHS) 2020 vision, a transformational, enterprise-wide approach to the health and human services business. HHS 2020 will move service delivery from a program-centric approach to a person-centric approach. New Mexico HSD will migrate away from program and technology silos into an integrated, flexible framework that supports service delivery and stakeholder interaction across HHS programs and organizations. HHS 2020 is technology-enabled, but includes rethinking organizational design, redesigning and streamlining business processes and reducing barriers between organizations within the HHS enterprise.

The HHS 2020 framework will be implemented through the MMISR project, the purpose of which is to replace the State's existing Medicaid Management Information System. HSD awarded Contract Number PSC 19-630-4000-0003 for International Business Machines Corporation (IBM) to implement the DS module as part of the overarching MMISR project and toward fulfilling the HHS 2020 Vision.

The DS module is responsible for designing, implementing, operating, and improving the structures, processes, and data needed to support HSD's and HHS 2020's current and future reporting and analytic requirements. The DS module for New Mexico will include an advanced data warehouse (DW) with business intelligence (BI), reporting, and other analytical capabilities. The DS module will be incrementally built out over nine contractually defined Design, Development, and Implementation (DDI) releases. This iterative approach allows IBM to deliver project features on a rolling basis. In addition, this incremental approach provides HSD with access to features and capabilities for testing and evaluation as they are completed.

For a detailed MMISR project overview, please see Section 1: Introduction in HSD's MMISR Project Management Plan.

1.3 Document Scope

The purpose of the Data Models deliverable is to provide IBM's approach to data modeling for the DS module, as well as the conceptual, logical, and physical data models for each data domain. This document will be updated during DDI, per the contract. DS will develop data models for each of the data domains as the SI's data models for each of the data domains becomes available.

The deliverable includes the following sections:

- **Introduction** – Introduces the deliverable and provides a project overview, document scope, and document maintenance
- **Assumptions, Constraints, and Risks** – Documents known assumptions, constraints, and risks related to the deliverable. A complete list of risks related to the project is maintained in the Risk Management Log Addendum
- **Requirements Traceability** – Documents the requirements pertaining to or fulfilled by the deliverable per the contract statement of work, HSD Request for Proposal (RFP), and IBM proposal responses
- **CMS Certification** – Documents how the deliverable relates to CMS Certification, and specifically, the Medicaid Enterprise Certification Toolkit (MECT) Checklist
- **Data Modeling Overview** – Documents the DS data modeling approach, methodologies, and roles and responsibilities
- **Conceptual Data Model –** Provides an overview, as well as access to the conceptual data model provided on HSD SharePoint
- **Logical Data Model -** Provides an overview, as well as access to the logical data model provided on HSD SharePoint
- **Physical Data Model -** Provides an overview, as well as access to the physical data model provided on HSD SharePoint

1.4 Definitions

Key terminology and acronyms needed for understanding the deliverable are defined in Appendix A.

1.5 Project Standards

This section describes standards of project management, communication, and security to be applied across the DS Data Models deliverable.

1.5.1 Project Management

It is the expectation that project standards, processes, and protocols established in the project initiation phase will continue through the implementation of the DS module. Any project management activities described in this document will comply with the HSD Project Management Office (PMO) Project Management Plan and Project Scheduling Plan, unless otherwise noted. This includes standards regarding project roles, responsibilities, scheduling, risk management, reporting, as well as all other project management activities.

1.5.2 Project Communications

Success of the project depends on strong and frequent communication conducted in an open and cooperative manner. The DS module will comply with the HSD PMO Communications Management Plan. The IBM team members will work with HSD, the SI contractor, and other stakeholders to ensure dependencies are communicated along with risks, issues, decisions, changes, and status.

Any project faces unexpected issues and problems. IBM will work with HSD, the SI contractor, and other MMISR partners to track and resolve issues and problems in a timely and efficient manner. Escalations will be used appropriately and documented in the DS Risk Management Log Addendum.

1.5.3 Security

IBM implements and maintains all security and privacy controls as required by Minimum Acceptable Risk Standards for Exchange (MARS-E) v2.0 standards. The implementation of these controls is documented in DS Module System Security Plan (SSP), accepted by HSD in February 2020. The SSP is a living plan that will be updated as the DS module progresses through DDI and into Operations and Maintenance (O&M). Once into O&M, the SSP will be reviewed and revised as necessary – annually at a minimum.

1.5.4 Document Maintenance

This deliverable will be reviewed by HSD and finalized according to the requirements defined in Section 3: Requirements Traceability. This document contains a Revision History log on the Document Information page. When changes occur, the version number will be updated to the next increment as well as its revision date and change description.

2 ASSUMPTIONS, CONSTRAINTS, AND RISKS

2.1 Assumptions

- The DS module will receive the enterprise data dictionary and model from the SI. The DS module is dependent on the accuracy of these documents
- DS data models will be iteratively developed based on data domain, as SI data domain data models and data are made available to DS
- The SI will be the primary data source for DS module data. This will include data received from the SI's System Migration Repository (SMR), operational data store (ODS), or any future modules
- HSD has identified the needed resources and subject matter experts (SMEs), and they will be available to attend associated working sessions and meetings as appropriate

2.2 Constraints

- DS can only provide data models for the data domains which the SI has provided data models and data for
- The Information Governance Catalog (IGC) is constrained by what DataStage can access in order to populate the IGC

2.3 Risks

- There are no risks as of this release.

3 REQUIREMENTS TRACEABILITY

This section provides the deliverable requirements as articulated using the following three sources:

- Contract Citation
- HSD RFP
- IBM RFP Response

3.1 Contract Citation

Table 1: Exhibit A, Section IV., D. Deliverable 04 Description

Description	Related Section
The Contractor shall create the DS Conceptual Data Model which identifies the highest-level data constructs and relationships between different business entities and is developed consistent with the nature of associated business processes. The DS Data Model - Conceptual will be developed in compliance with Enterprise Data Architecture (EDA) standards. The Conceptual Data Model will include the following: 1. Definition of the major entities of interest in terms that are meaningful to the way the HHS programs conducts MITA business processes. 2. Definition of a high-level relational map of the subject areas and cross-subject area dependencies for the data warehouse. 3. Capturing of these subject area relationships.	6
The Contractor shall create the DS Logical Data Model which is a representation of business concepts laid out in visual form that clearly shows these concepts and their various relationships. The Conceptual Data Model will provide the foundation for the Logical Data Model. The Logical Data Model will be developed in compliance with EDA standards. The Contractor shall deliver the Logical Data Model in XML Interchange (XMI) format. The Contractor shall use CA ERwin to build the Logical Data Model for the Data Services Module. The DS Logical Data Model shall include entities (tables), attributes (columns/fields) and relationships (keys). It will use business names for entities and attributes. It will be a data model of a specific domain whose expression is independent of a particular database management product or storage technology (platform/RDBMS) but is expressed in terms of data structures such as relational tables and columns.	7

Description	Related Section
The Logical Data Model shall be independent of the underlying physical implementation. The Logical Data Model will include the following features: 1. Each entity in the logical data model will be assigned a primary key—the attribute or set of attributes that distinguish one instance of the entity from another. 2. All the attributes for each entity will be included. Relationships between entities will be represented through foreign keys associated with the primary keys of the referring entity. 3. Normalization decisions will be finalized in the logical data model, which will result in the final normalized representation of entity-to-entity relationships.	
The Contractor shall develop the DS Physical Data Model such that it represents how the model will be built in the database. A physical database model shall show all table structures, including column name, column data type, column constraints (i.e., validation rules), primary key, foreign key, database triggers, stored procedures, domains, access constraints, indices for performance, and relationships between tables. The physical data model shall maximize the features of the RDBMS. The Physical Data Model shall be developed in compliance with EDA standards and shall be delivered in XML Interchange (XMI) format. The Physical Data Model shall include the following: 1. Specifications for all tables and columns. 2. Foreign keys used to identify relationships between tables. 3. De-normalization based on user requirements and performance considerations. 4. Physical considerations causing the physical data model to be different from the logical data model.	8

3.2 HSD RFP Requirements

The RFP-indicated requirements for the Release Strategy and subsequent IBM proposed response are described in the following table.

Table 2: RFP Appendix H Requirement 1.01 Description and Response

Objective Category	Requirement Number	Requirement	Related Section
System Design and Architecture	3.29	Offeror shall describe how its proposed Solution will optimize the retrieval and utilization of multi-dimensional data.	5.1.4

Objective Category	Requirement Number	Requirement	Related Section
System Design and Architecture	3.42	Offeror shall describe how its proposed Solution will import and maintain State-identified reference data bases (e.g., Census or Medicare data, commercial insurance data sets, etc.).	5.1.5
System Design and Architecture	3.45	Offeror shall describe how its proposed Solution will include tools that establish and manage a logical data model, and will include standards, responsibilities, relationships, definitions, domains, keys, and entity-relationship diagrams (ERDs).	5.1.1
System Design and Architecture	3.47	Offeror shall describe how its proposed Solution will inform/extend the data models in the DS and SI Architecture, including metadata management, to minimize maintenance tasks when additional data elements are produced or required by the Enterprise.	5.3.1
System Design and Architecture	3.48	Offeror shall describe how its proposed Solution will have the capability to update or extend approved schemas/models to incorporate new data fields as needed or requested.	5.2.4
System Design and Architecture	3.49	Offeror shall describe how its proposed Solution will support physical-to-logical model mapping and rationalization of its data translation, data aggregation and data augmentation processes, and provide definition of model-to-model relationships of repository objects, data aggregation and flows utilizing graphical attribute-level mapping.	5.1.1
System Design and Architecture	3.51	Offeror shall describe how its proposed Solution will minimize maintenance tasks when additional data elements are produced or required by the Enterprise.	5.2.6
Visualization, Reports and Analytics	9.03	Offeror shall describe how its proposed Solution will update report data fields as needed or requested and will provide for New Mexico-specific and user-defined report data fields for Enterprise use which will be revisable and expandable over time as necessary to support Enterprise program requirements.	5.2.1

3.3 IBM Proposal Responses

The IBM proposal responses to the above HSD RFP requirements are contained in Appendix B.

4 CMS CERTIFICATION

The DS Data Models deliverable is linked to the following CMS Certification requirements, and specifically the MECT Checklist.

Table 3: CMS Certification

MECT Business Process:	Data Management Strategy (DMS)
MECT Business Area:	Information Architecture
Criteria ID:	IA.DMS.5
Criteria:	The system refreshes or replaces all historical claim data, recipient enrollment, provider enrollment, and other primary reference data on a scheduled basis.
Deliverable Section Reference:	Logical Data Model

MECT Business Process:	Logical Data Model (LDM)
MECT Business Area:	Information Architecture
Criteria ID:	IA.LDM.5
Criteria:	The system of interest Logical Data Model (LDM) supports identification of data classes, attributes, relationships, standards, and code sets for intrastate exchange.
Deliverable Section Reference:	Logical Data Model

5 DATA MODELING OVERVIEW

- **Section 5.1: Approach** – Details the tools and schema layout, as well as Extract, Transform, and Load (ETL) processing, data retrieval, and reference data bases
- **Section 5.2: Schema Layout and Data Flow** – Displays the schema layout for the DS module
- **Section 5.3: Modeling Methodologies** – Documents the DS data modeling methods, including data fields, types, data vault, star schemata, and maintenance
- **Section 5.4: Standards** – Details the enterprise-wide standards defined by HSD and the SI to which DS data modeling will comply
- **Section 5.5: Roles and Responsibilities** – Defines the IBM and HSD roles responsible for data modeling

5.1 Approach

The DS module will produce data models consistent with the data models provided by the SI contractor. The development of DS data models will follow the order and cadence of SI delivery of data models by data domain and the Hybrid Agile iterative development approach outlined in the DS contract.

As data models for each data domain become available, DS will iteratively develop the relevant data models over the course of DDI. The data domains are detailed in the DS System Design Document deliverable. As of DDI Release 4, DS has developed data models for all domains.

Note: The linked data models for the MCO domain are subject to change based on expected upcoming changes to the data models from the SI.

The detailed approach to DS data model development is provided over the following subsections.

5.1.1 Data Modeling Tools

RFP 3.45: Offeror shall describe how its proposed Solution will include tools that establish and manage a logical data model, and will include standards, responsibilities, relationships, definitions, domains, keys and entity-relationship diagrams (ERDs).

The following table details the tools used for DS data modeling and data warehousing, including their use.

Table 4: Tools

Tool Name	Use
Cognos	Framework Manager provides modeling of the three semantic layers: data layer, business layer, and presentation layer
DataStage	ETL tool used for data warehousing

ERwin	Data modeling tool used to create conceptual, logical, and physical data models
Information Governance Catalog	Used to display data modeling and data transformation from ERwin, DataStage, and Cognos to the end user
Oracle Structured Query Language (SQL) Developer	SQL client used to navigate and query schemata and tables in the Oracle database
Service Now	Web service ticketing application used to assign tasks to the DBA team for creating tables and other database objects as required by the data models
Sparx Enterprise Architect / Sparx Cloud	Modeling software and server for central storage of data models
Tableau	Provides modeling for metrics, key performance indicators (KPIs), and other calculated fields

5.1.2 Rationalization

RFP 3.49: Offeror shall describe how its proposed Solution will support physical-to-logical model mapping and rationalization of its data translation, data aggregation and data augmentation processes, and provide definition of model-to-model relationships of repository objects, data aggregation and flows utilizing graphical attribute-level mapping.

DS does all physical, logical, and conceptual data mapping within the ERwin data modeling tool. Cognos and DataStage will perform data aggregation and flows, and these flows are then used by the IGC tool. Additionally, there is an interface to the IGC that allows these data models to be viewed. Through that functionality, physical-to-logical model mapping and rationalization will be provided by the IGC tool.

5.2 Schema Layout and Data Flow

The following table displays the schema layout for the DS module, including definitions and associated model method. These methods are further described in subsequent sections.

Table 5: Schema Names, Descriptions, and Model Method

Schema Name	Definition	Model Method
Source Schema (SRC)_XXX	Source schema that holds incoming data from a specific system, designated here as XXX	Type 1, truncate and refresh

Schema Name	Definition	Model Method
	All the columns in SRC schema is string (VARCHAR) to accommodate all the incoming data	
Staging Schema (STG)_XXX	Staging schema that adds columns used to manage warehouse processes and models The staging schema prepares the data for data warehouse (Integrated [INT] schema) by adding data warehouse related columns such as ETL metadata columns, hash keys etc., and performs quality checks on the data, such as data type validation and uniqueness based on the data model	Type 1, truncate and refresh
Persistent Storage Area (PSA)_ _XXX	Persistent storage area schema that holds a copy of all data that has arrived in the source schema All the columns in PSA schema is string (VARCHAR) to accommodate all the incoming data	Storage area that is a copy of source schema with additional batch ID column. This is an insert only or history table
Integrated Schema (INT)	Integrated schema that contains all source data from all source systems	Data vault 2 and Type 2, slowly changing dimensions
Enterprise Event Model Schema (EEM)	Enterprise event model schema that is optimized for data querying	Star schema
Mart_XXX	Data mart schema optimized for data querying specific to a department or group	Star schema
Reports and Extracts Schema (REM)	Schema optimized for extracts and canned reports	Type 1, truncate and refresh

5.2.1 ETL Processing

ETL processing allows data to move through the DS system. The data will first load into the SRC_XXX schema. A copy of the source data will be placed in the PSA_XXX schema. From there, the data will be

loaded to the STG_XXX schema to be reconciled. This will involve confirming counts, valid values, and standardized data types. The process for loading the data will be specific to the incoming table. This allows parallel processing to occur, thereby eliminating wait times or dependencies incurred by serial processing.

After entering successfully through STG_XXX schema, the data will be loaded into the warehouse. To help speed the process of loading data, foreign key (FK) relationships between tables will be enforced programmatically instead of by foreign key indexes. This permits the loading to continue even if dependent parent records have not been delivered to the DS module in a timely fashion.

5.2.2 Data Retrieval and Utilization

RFP 3.29: Offeror shall describe how its proposed Solution will optimize the retrieval and utilization of multi-dimensional data.

The data marts and the EEM schema will have data models based upon the star schema and cube methodologies. Cognos and Tableau can both utilize the star schema and cube structures to optimize retrieval for report creation and publication.

5.2.3 Reference Data Bases

RFP 3.42: Offeror shall describe how its proposed Solution will import and maintain State-identified reference data bases (e.g., Census or Medicare data, commercial insurance data sets, etc.).

If through the SI, reference data checks will be loaded according to the normal process described in previous sections. If outside of the SI, HSD will need to provide a data dictionary of the incoming tables along with an ERD of the incoming tables. This would then be placed on the Secure File Transfer Protocol (SFTP) server and ingested using the methodology outlined in the subsequent sections.

5.3 Modeling Methodologies

The following warehousing methodologies are used within the schema layouts and data flow described in the above section.

5.3.1 Data Fields

RFP 9.03: Offeror shall describe how its proposed Solution will update report data fields as needed or requested and will provide for New Mexico-specific and user-defined report data fields for Enterprise use which will be revisable and expandable over time as necessary to support Enterprise program requirements.

Data fields are the columns within a table. They have specific data types, such as text, numeric, binary, and Boolean. The content of the data fields may also be constrained by values in a reference table. In addition, business rules can constrain data field values, e.g. ages between 0 and 125. DS data fields will enforce the canonical model data types and values provided by the SI contractor.

Report data fields are by default derived from the enterprise canonical model data fields. When an update happens to the enterprise canonical model data fields, the DS data architect will update the data models across all schemata. When Cognos is opened, it rereads the underlying tables that are available for the user. If the column names did not change, but the data type did change, the DS report data fields will be automatically updated. If the column names have been changed or new columns added that are used within the report, the report creator will have to add the columns to the report.

For report data fields modified or not previously in the canonical model, impact analysis from changes to the canonical model can be performed using the IGC to identify report data fields that must be updated manually.

5.3.2 Type 1

The Type 1 data warehousing methodology involves emptying a table and then repopulating a table upon a new load. This is also known as a truncate and refresh approach.

5.3.3 Type 2

The Type 2 data warehousing methodology involves adding a Load Date and Last Update date column, as well as a hash key and a hash diff key, to all incoming data.

Incoming records will be identified by their primary key and assigned a hash key value using the secure hash algorithm 1 (SHA-1). The remaining columns will also be given a hash value called hash diff key.

1. If the hash key value does not exist in the current table, the record will be inserted with the current date and time in the Load Date and Last Update columns. The record with the maximum Last Update column is the most current record for this hash key combination record
2. If the hash key value does exist in the current table, the hash diff key values will be compared to determine if there is any difference
3. If the hash key diff values are the same, the new record is ignored. It is not entered as new data
4. If there are differences, a new record is inserted with the initial load date and the current date and time inserted into the Last Update column
5. Additionally, another column is added in the EEM schema with a current record indicator for performance reasons only

5.3.4 Data Vault

RFP 3.48: Offeror shall describe how its proposed Solution will have the capability to update or extend approved schemas/models to incorporate new data fields as needed or requested.

The data vault model methodology is flexible, scalable, consistent, and highly adaptable to any source system. Data objects are modeled using hub, link, and satellite concepts, which are defined as follows:

1. **Hubs** - Uniquely identify business entities, such as provider and client
2. **Links** - Record the transactions between the hubs. An example of a link record is a client's office visit to a provider

3. **Satellites -** Give detail about hub or link records and are system specific. Demographic data from the SMR associated with a client would be an example of a satellite record

Since satellites are system specific, any changes from the source system to that specific satellite minimizes impact on the overall data system. Because this methodology is more resilient than other current industry modeling techniques, the DS module is proficient at maintaining integrity.

5.3.5 Star Schemata

A star schema is a type of data model that is optimized for removing data from the system, which is why it is used for reporting. Star schemata are used within the EEM and specific data marts.

At the center of the star schema is a transaction table, also known as a fact table. It contains numeric data about the transaction and keys to description data tables, known as dimensions. A dimension is a description table that is typically hierarchical. As an example, a fact table may have a date key that references the specific record in a date dimension. The date dimension would have a hierarchy for a date such as month, quarter, and year. Utilizing this data model, a query can be performed to find all transactions within a specific quarter.

5.3.6 Maintenance

RFP 3.51: Offeror shall describe how its proposed Solution will minimize maintenance tasks when additional data elements are produced or required by the Enterprise.

Since DS is using data vault methodology, source systems have their own data structures (satellites). Therefore, when a change occurs, only the source system data structure needs to be modified. Cognos and Tableau automatically refresh semantic layers from the data marts as the data marts are updated, either with additional data elements or structures.

5.4 Standards

DS will implement enterprise wide standards, as well as specific modeling standards, as detailed in the following sections.

5.4.1 Enterprise Standards

RFP 3.47: Offeror shall describe how its proposed Solution will inform/extend the data models in the DS and SI Architecture, including metadata management, to minimize maintenance tasks when additional data elements are produced or required by the Enterprise.

In order for data to be highly usable within HSD and across the other New Mexico State agencies, it should comply to agreed-upon standards. DS will use the standards provided in the SI's enterprise data model and data dictionary. In addition, the SI's SMR will be compliant with Federal Health Information Requirements (FHIR), Federal Health Information Model (FHIM), and National Information Exchange Model (NIEM) per the HSD Data Management Workgroup. As a result, DS will load and use FHIR-, FHIM-, and NIEM-compliant data.

5.4.2 Modeling Standards

In addition to the enterprise-wide standards, DS data models will be developed in compliance with EDA standards. Data models will be built using CA ERwin. For ease of HSD and CMS access and use, data models will be provided as Portable Document Format (PDF) files on the HSD SharePoint site and linked to this deliverable in the appropriate section. Data models will be delivered in the agreed format of XMI upon request. The XMLs created are placed in the HSD SharePoint site.

The files in the Data Models folder are enumerated in the following sections:

- Section 6: Conceptual Data Model
- Section 7: Logical Data Model
- Section 8: Physical Data Model

5.5 Roles and Responsibilities

5.5.1 IBM Roles and Responsibilities

The roles and responsibilities for IBM personnel pertaining to the DS Data Models deliverable are presented in the table below. For the complete record of the staff on the project team, as well as the process and procedures used to manage staff throughout the project's life, please refer to the Resource Management Plan.

Table 6: IBM Roles and Responsibilities

IBM Role	IBM Responsibilities
BI Architect	• Creates star schemata in EEM schema and other marts as needed
Data Architect	• Creates tables in SRC, PSA, and INT that adhere to the data model methodologies outlined in this document • Creates tables and columns that persist the naming of the data elements throughout the DS module as appropriate • Provides accurate mappings to the ETL team for loading the data into these appropriate data structures
ETL Team	• Loads data into the database and to all schemata
Quality Assurance (QA) Team	• Tests the data models were accurately created

5.5.2 HSD Project Responsibilities

The HSD leadership team roles and responsibilities related to the DS Data Models deliverable are presented in the table below.

Table 7: HSD Roles and Responsibilities

HSD Role	HSD Responsibilities
Enterprise Data Architect	• Verifies that the SI data models supplied to the DS module comply with the Enterprise Data Model and HSD Data Standards
QA Testing Team	• Verifies that the IGC fulfills the requirements as specified in the contract and RFP
Primary Business Lead	• Provides feedback to the development team on the metadata catalog functionality of the IGC

6 CONCEPTUAL DATA MODEL

Contract: The Contractor shall create the DS Conceptual Data Model which identifies the highest-level data constructs and relationships between different business entities and is developed consistent with the nature of associated business processes. The DS Data Model - Conceptual will be developed in compliance with Enterprise Data Architecture (EDA) standards.

The Conceptual Data Model will include the following:

1. *Definition of the major entities of interest in terms that are meaningful to the way the HHS programs conducts MITA business processes.*
2. *Definition of a high-level relational map of the subject areas and cross-subject area dependencies for the data warehouse.*
3. *Capturing of these subject area relationships.*

6.1 Overview

The Conceptual Data Model identifies the highest-level relationships between major entities of interest in terms of how HSD conducts MITA business processes. The Conceptual Data Model will be iteratively developed over the course of DDI as the data models and data relative to each data domain becomes available from the SI.

Please find entity names and descriptions stored on HSD SharePoint here.

The DS module has the following domains:

- Claim
- Client
- Drug Rebate
- Finance
- Managed Care Organization (MCO)
- Prior Authorization
- Provider
- Third-Party Liability (TPL)

The major entities are at the table level, as dictated by the SI.

The relationships between subject areas as applicable to the current level of data model and data detail available are in the DS Physical Data Model provided in Section 8.

6.2 Features of the Conceptual Data Model

The following describes the general features of a conceptual data model:

- Coverage of the business concepts

- Designed and developed primarily for a business audience
- Contains relationships between entities
- Designed and developed to be independent of Database Management System (DBMS), data storage locations, or technologies

6.3 Diagrams

Conceptual models were generated based on those provided by the SI. Data warehousing schemata, such as INT and EEM, do not use conceptual models.

The following conceptual data models are provided on HSD SharePoint and can be found at the following links.

Table 8: Related Schemata and Data Models - Conceptual

Related Schema	Related Data Model	Comment
SRC, STG, and PSA	• Client ◊ Client_Conceptual_Model_Part1 ◊ Client_Conceptual_Model_Part2 ◊ Client_Conceptual_Model_Part3 ◊ Client_Conceptual_Model_Part4 ◊ Client_Conceptual_Model_Part5 • Drug Rebate ◊ Drug_Rebate_Conceptual_Model • Finance ◊ Finance_Conceptual_Model • MCO ◊ Managed_Care_Organization_Conceptual_Model • Prior Authorization ◊ Prior Authorization Conceptual Model • Provider ◊ Provider_Authorization Conceptual_Model • TPL ◊ Third_Party Liability_Conceptual_Model • Claim ◊ Claim Conceptual Model Part 1 ◊ Claim Conceptual Model Part 2 ◊ Claim Conceptual Model Part 3 ◊ Claim Conceptual Model Part 4 ◊ Claim Conceptual Model Part 5	**Note:** For legibility, the data model for the Client domain and the data model for the Claim domain were each divided into five parts.

7 LOGICAL DATA MODEL

The Contractor shall create the DS Logical Data Model which is a representation of business concepts laid out in visual form that clearly shows these concepts and their various relationships. The Conceptual Data Model will provide the foundation for the Logical Data Model. The Logical Data Model will be developed in compliance with EDA standards. The Contractor shall deliver the Logical Data Model in XML Interchange (XMI) format. The Contractor shall use CA ERwin to build the Logical Data Model for the Data Services Module. The DS Logical Data Model shall include entities (tables), attributes (columns/fields) and relationships (keys). It will use business names for entities and attributes. It will be a data model of a specific domain whose expression is independent of a particular database management product or storage technology (platform/RDBMS) but is expressed in terms of data structures such as relational tables and columns. The Logical Data Model shall be independent of the underlying physical implementation. The Logical Data Model will include the following features:

1. *Each entity in the logical data model will be assigned a primary key—the attribute or set of attributes that distinguish one instance of the entity from another.*
2. *All the attributes for each entity will be included. Relationships between entities will be represented through foreign keys associated with the primary keys of the referring entity.*
3. *Normalization decisions will be finalized in the logical data model, which will result in the final normalized representation of entity-to-entity relationships.*

7.1 Overview

The Logical Data Model is a representation of business concepts and relationships. It is independent of the underlying physical implementation. The Logical Data Model will be iteratively developed over the course of DDI as the data models and data relative to each data domain becomes available from the SI. The DS module domains will provide the source data for the Flexible Analytics (FA) service, which in turn populates the FA data mart (see **Error! Reference source not found.** below).

The DS module has the following:

Domains

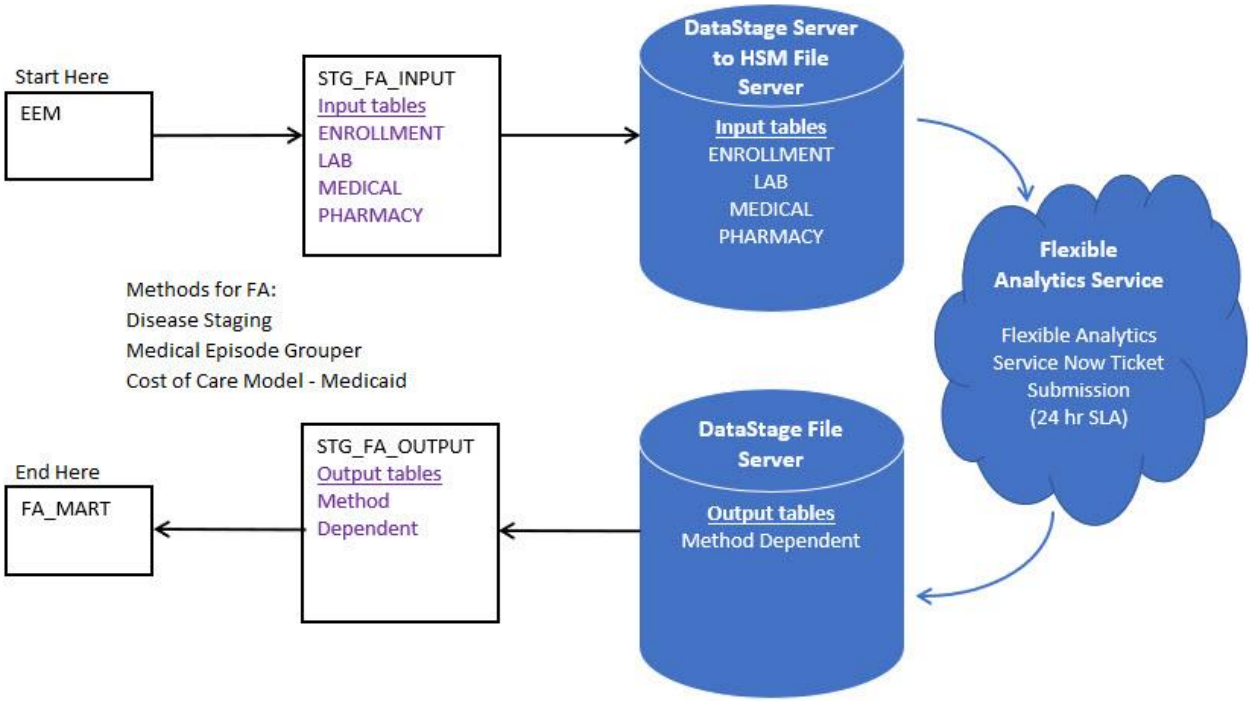
- Claim
- Client
- Drug Rebate
- Finance
- MCO
- Prior Authorization
- Provider
- TPL

Analytics

- FA

The figure below shows the data flow between the DS module schema and the FA service.

Figure 1: FA Data Flow



7.2 Features of the Logical Data Model

Per the contract, the following features are of note in the Logical Data Model:

- Primary keys are indicated on the model by use of a key symbol
- Attributes, also known as column names, are listed within each entity
- Because DS inherits the data model from the SI, normalization has not been applied to these models at the DS level. If normalization occurred, it was performed by the SI

7.3 Diagrams

The logical data models are provided on HSD SharePoint and can be found at the links in the following table.

Table 9: Related Schemata and Data Models - Logical

Related Schema	Related Data Model	Comments
SRC, PSA, STG	• Client ◊ Client Source PSA Model Part 1 ◊ Client Source PSA Model Part 2	**Note:** For legibility, the data models for

Related Schema	Related Data Model	Comments
	◊ Client Source PSA Model Part 3 ◊ Client Source PSA Model Part 4 ◊ Client Source PSA Model Part 5 ◊ Client Stage Model Part 1 ◊ Client Stage Model Part 2 ◊ Client Stage Model Part 3 ◊ Client Stage Model Part 4 ◊ Client Stage Model Part 5 • Drug Rebate ◊ Drug Rebate Source PSA Model ◊ Drug Rebate Stage Model • Finance ◊ Finance Source PSA Model ◊ Finance Stage Model • MCO ◊ Managed_Care_Organization_Source_PSA_Model ◊ Managed_Care_Organization_Stage_LDM • Prior Authorization ◊ Prior Authorization SRC and PSA Schema ◊ Prior Authorization STG Schema • Provider ◊ Provider Source PSA Model ◊ Provider Stage Model • TPL ◊ Third Party Liability Source PSA Model ◊ Third Party Liability STG Data Model • Claim ◊ Claim_SRC_PSA_LDM_Part1 ◊ Claim_SRC_PSA_LDM_Part2 ◊ Claim_SRC_PSA_LDM_Part3 ◊ Claim_SRC_PSA_LDM_Part4 ◊ Claim_SRC_PSA_LDM_Part5 ◊ Claim_Staging_LDM_Part1 ◊ Claim_Staging_LDM_Part2 ◊ Claim_Staging_LDM_Part3 ◊ Claim_Staging_LDM_Part4 ◊ Claim_Staging_LDM_Part5	the Client domain and the data models for the Claim domain were each divided into five parts.

Related Schema	Related Data Model	Comments
INT	• Client ◊ Client INT (DataVault)_Model LDM Part 1 ◊ Client INT (DataVault)_Model LDM Part 2 ◊ Client INT (DataVault)_Model LDM Part 3 ◊ Client INT (DataVault)_Model LDM Part 4 ◊ Client INT (DataVault)_Model LDM Part 5 ◊ Client INT (DataVault)_Model LDM Part 6 ◊ Client INT (DataVault)_Model LDM Part 7 • Finance ◊ Finance INT (DataVault) Model LDM • Flexible Analytics ◊ FA_Admission_INT_LDM ◊ FA_Disease Staging_INT_LDM • MCO ◊ Managed_Care_Organization_INT_(DataVault)_Model_LDM • Prior Authorization ◊ Prior Authorization INT (DataVault) Schema ◊ Prior Authorization (DataVault) Model LDM • Provider ◊ Provider INT (DataVault) Model LDM • TPL ◊ Third Party Liability INT (DataVault) Model LDM • Claim ◊ Claim_INT_Model_LDM_Part1 ◊ Claim_INT_Model_LDM_Part2 ◊ Claim_INT_Model_LDM_Part3 ◊ Claim_INT_Model_LDM_Part4 ◊ Claim_INT_Model_LDM_Part5 ◊ Claim_INT_Model_LDM_Part6	**Note:** For legibility, the data models for the Client domain was divided into seven parts, and the data model for the Claim domain was divided into six parts.
EEM	• Client ◊ Client_EEM_Address_Event_Fact_LDM ◊ Client_EEM_Authorized_Representative_Address_Event_Fact_LDM ◊ Client_EEM_Copay_Event_Fact_LDM ◊ Client_EEM_Eligibility_Event_Fact_LDM ◊ Client_EEM_Enrollment LockIn_Event_Fact_LDM ◊ Client_EEM_Enrollment_Event_Fact_LDM ◊ Client_EEM_Event_Fact_LDM ◊ Client_EEM_Level_of_Care_Event_Fact_LDM	**Note:** For legibility, the dimensions data model for the Claim domain was divided into eight parts.

Related Schema	Related Data Model	Comments
	◊ Client_EEM_Long_Term_Care_Liability_Event_Fact_LDM ◊ Client_EEM_Medicare_Part_C_Enrollment_Event_Fact_LDM ◊ Client_EEM_Medicare_Part_D_Span_Event_Fact_LDM ◊ Client_EEM_Merge_Event_Fact_LDM ◊ Client_EEM_Pregnancy_Event_Fact_LDM ◊ Client_EEM_Relationship_Head_of_Household_Event_Fact_LDM • Finance ◊ Finance EEM Model LDM • MCO ◊ MCO_Allotment_Details_Event_Fact_LDM ◊ MCO_EEM_Cohort_Age_Event_Fact_LDM ◊ MCO_EEM_Cohort_Category_Of_Eligibility_Event_Fact_LDM ◊ MCO_EEM_Dimensions_LDM ◊ MCO_EEM_MCO_Detail_Event_Fact_LDM ◊ MCO_EEM_Plan_Detail_Event_Fact_LDM ◊ MCO_EEM_Plan_Rate_Event_Fact_LDM ◊ MCO_EEM_Plan_Specialty_Exclusion_Event_Fact_LDM • Prior Authorization ◊ Prior Authorization EEM Schema Detail ◊ Prior Authorization EEM Schema Header ◊ Prior Authorization Exception EEM Model ◊ Prior Authorization Detail EEM LDM ◊ Prior Authorization Exception EEM LDM ◊ Prior Authorization Header EEM LDM • Provider ◊ Provider_EEM_Dimensions_LDM ◊ Provider_EEM_Provider_Address_Event_Fact_LDM ◊ Provider_EEM_Provider_Clia_Certification_Event_Fact_LDM ◊ Provider_EEM_Provider_Disproportionate_Share_Event_Fact_LDM ◊ Provider_EEM_Provider_Electronic_Fund_Transfer_Event_Fact_LDM ◊ Provider_EEM_Provider_Enrollment_Event_Fact_LDM ◊ Provider_EEM_Provider_Event_Fact_LDM ◊ Provider_EEM_Provider_License_Certificate_Event_Fact_LDM ◊ Provider_EEM_Provider_Notes_Event_Fact_LDM ◊ Provider_EEM_Provider_Npi_Event_Fact_LDM	

Related Schema	Related Data Model	Comments
	◊ Provider_EEM_Provider_Owner_Employee_Event_Fact_LDM ◊ Provider_EEM_Provider_Review_Claim_Event_Fact_LDM ◊ Provider_EEM_Provider_Review_Program_Event_Fact_LDM ◊ Provider_EEM_Provider_Taxonomy_Event_Fact_LDM • TPL ◊ Third Party Liability EEM LDM • Claim ◊ Claim_EEM_Model_Detail_Dental_Event_Fact_LDM ◊ Claim_EEM_Model_Detail_Professional_Event_Fact_LDM ◊ Claim_EEM_Model_Diagnosis_Event_Fact_LDM ◊ Claim_EEM_Model_Dimensions_LDM_Part1 ◊ Claim_EEM_Model_Dimensions_LDM_Part2 ◊ Claim_EEM_Model_Dimensions_LDM_Part3 ◊ Claim_EEM_Model_Dimensions_LDM_Part4 ◊ Claim_EEM_Model_Dimensions_LDM_Part5 ◊ Claim_EEM_Model_Dimensions_LDM_Part6 ◊ Claim_EEM_Model_Dimensions_LDM_Part7 ◊ Claim_EEM_Model_Dimensions_LDM_Part8 ◊ Claim_EEM_Model_Exception_Event_Fact_LDM ◊ Claim_EEM_Model_Financial_Event_Fact_LDM ◊ Claim_EEM_Model_Header_Dental_Event_Fact_LDM ◊ Claim_EEM_Model_Header_Professional_Event_Fact_LDM ◊ Claim_EEM_Model_Lineitem_COB_Adjustment_LDM ◊ Claim_EEM_Model_MSQ_Event_Fact_LDM ◊ Claim_EEM_Model_Override_Exception_Event_Fact_LDM ◊ Claim_EEM_Model_Previous_Location_Event_Fact_LDM ◊ Claim_EEM_Model_Provider_Event_Fact_LDM ◊ Claim_EEM_Model_Surgical_Procedure_Event_Fact_LDM	
STG_FA_INPUT	• Flexible Analytics – Input ◊ FA_Admission_STG_FA_INPUT_LDM ◊ FA_Disease Staging_STG_FA_INPUT_LDM ◊ FA_Stage_INPUT_LDM	
STG_FA_OUTPUT	• Flexible Analytics – Output ◊ FA_Admission_STG_FA_OUTPUT_LDM ◊ FA_Disease Staging_STG_FA_OUTPUT_LDM	
FA_MART	• Flexible Analytics ◊ FA_Admission_MART_LDM	

Related Schema	Related Data Model	Comments
	◊ FA_Disease Staging_MART_LDM	

8 PHYSICAL DATA MODEL

The Contractor shall develop the DS Physical Data Model such that it represents how the model will be built in the database. A physical database model shall show all table structures, including column name, column data type, column constraints (i.e., validation rules), primary key, foreign key, database triggers, stored procedures, domains, access constraints, indices for performance, and relationships between tables. The physical data model shall maximize the features of the RDBMS. The Physical Data Model shall be developed in compliance with EDA standards and shall be delivered in XML Interchange (XMI) format.

The Physical Data Model shall include the following:

1. *Specifications for all tables and columns*
2. *Foreign keys used to identify relationships between tables*
3. *De-normalization based on user requirements and performance considerations*
4. *Physical considerations causing the physical data model to be different from the logical data model*

8.1 Overview

The Physical Data Model represents how the model will be built in the database. A physical database model shows all table structures, including column name, column data type, column constraints, primary key, foreign key, and relationships between tables. The Physical Data Model will be iteratively developed over the course of DDI as the data models and data relative to each data domain becomes available from the SI. The DS module domains will provide the source data for the FA service, which in turn populates the FA data mart (see **Error! Reference source not found.**).

The DS module has the following:

Domains

- Claim
- Client
- Drug Rebate
- Finance
- MCO
- Prior Authorization
- Provider
- TPL

Analytics

- FA

The physical ERD represents the actual design of the model with respect to the database. It represents how data should be structured and related in a specific database, so it is important to consider the

naming conventions, defining referential integrity, defining primary key(s) and applying the restrictions for a database management system, if any.

The following list provides the major differences of a physical model from a logical model:

- Entity names are now table names
- Attributes are now column names
- Data type for each column is specified. Data types can be different depending on the actual database being used
- Adding primary keys, foreign keys, and constraints to the design that reflect the actual physical column names
- Denormalization may occur depending on the methodology (data vault or dimensional model or 3rd normal form)
- Many-to-many relationships are modeled

Per the contract, the following features are of note in the Physical Data Model:

- Specifications for all tables and columns are provided in the physical model itself
- Foreign keys are identified in the model using FK
- De-normalization was not performed at this time but may be performed in later releases
- During ingestion (SRC, PSA, STG), DS must accommodate the physical data model set by the SI. DS has logical models for the INT and EEM schemata, which are different models from the SI

8.2 Diagrams

The physical data models are provided on HSD SharePoint and can be found at the links in the following table:

Table 10: Related Schemata and Data Models - Physical

Related Schema	Related Data Model	Comment
SRC, STG, PSA	• Client ◊ Client Source PSA Model Part 1 ◊ Client Source PSA Model Part 2 ◊ Client Source PSA Model Part 3 ◊ Client Source PSA Model Part 4 ◊ Client Source PSA Model Part 5 ◊ Client_Stage_Model_Part1 ◊ Client_Stage_Model_Part2 ◊ Client_Stage_Model_Part3 ◊ Client_Stage_Model_Part4 ◊ Client_Stage_Model_Part5 ◊ Client_Stage_Model_Part6	**Note:** For legibility, the data models for the Client domain were divided into five and six parts, and the data models for the Claim domain were each divided into five parts.

Related Schema	Related Data Model	Comment
	• Drug Rebate ◊ Drug Rebate Source PSA Model ◊ Drug Rebate Stage Model • Finance ◊ Finance Source PSA Model ◊ Finance Stage Model • MCO ◊ Managed_Care_Organization_Source_PSA_Model ◊ Managed_Care_Organization_Stage_PDM • Prior Authorization ◊ Prior Authorization SRC and PSA Schema ◊ Prior Authorization STG Schema • Provider ◊ Provider Source PSA Model ◊ Provider Stage Model • TPL ◊ Third Party Liability Source PSA Model ◊ Third Party Liability STG Data Model • Claim ◊ Claim_SRC_PSA_Model_Part1 ◊ Claim_SRC_PSA_Model_Part2 ◊ Claim_SRC_PSA_Model_Part3 ◊ Claim_SRC_PSA_Model_Part4 ◊ Claim_SRC_PSA_Model_Part5 ◊ Claim_Staging_PDM_Part1 ◊ Claim_Staging_PDM_Part2 ◊ Claim_Staging_PDM_Part3 ◊ Claim_Staging_PDM_Part4 ◊ Claim_Staging_PDM_Part5	
INT	• Client ◊ Client INT (DataVault) Model PDM Part 1 ◊ Client INT (DataVault) Model PDM Part 2 ◊ Client INT (DataVault) Model PDM Part 3 ◊ Client INT (DataVault) Model PDM Part 4 ◊ Client INT (DataVault) Model PDM Part 5 ◊ Client INT (DataVault) Model PDM Part 6 ◊ Client INT (DataVault) Model PDM Part 7 • Finance ◊ Finance INT (DataVault) Model PDM	**Note:** For legibility, the data model for the Client domain was divided into seven parts, and the data model for the Claim domain was

Related Schema	Related Data Model	Comment
	• Flexible Analytics ◊ FA_Admission_INT_PDM ◊ FA_Disease Staging_INT_PDM • MCO ◊ Managed_Care_Organization_INT_(DataVault)_Model_PDM • Prior Authorization ◊ Prior Authorization INT Schema • Provider ◊ Provider INT (DataVault) Model PDM • TPL ◊ Third Party Liability INT (DataVault) Model PDM • **TPL – DDI Release 7** ◊ TPL INT Model_Logical ◊ TPL INT Model_Physical • Claim ◊ Claim_INT_Model_PDM_Part1 ◊ Claim_INT_Model_PDM_Part2 ◊ Claim_INT_Model_PDM_Part3 ◊ Claim_INT_Model_PDM_Part4 ◊ Claim_INT_Model_PDM_Part5 ◊ Claim_INT_Model_PDM_Part6	divided into six parts.
EEM	• Claim ◊ Claim_EEM_Model_Detail_Dental_Event_Fact_PDM ◊ Claim_EEM_Model_Detail_Professional_Event_Fact_PDM ◊ Claim_EEM_Model_Diagnosis_Event_Fact_PDM ◊ Claim_EEM_Model_Dimensions_PDM_Part1 ◊ Claim_EEM_Model_Dimensions_PDM_Part2 ◊ Claim_EEM_Model_Dimensions_PDM_Part3 ◊ Claim_EEM_Model_Dimensions_PDM_Part4 ◊ Claim_EEM_Model_Dimensions_PDM_Part5 ◊ Claim_EEM_Model_Dimensions_PDM_Part6 ◊ Claim_EEM_Model_Dimensions_PDM_Part7 ◊ Claim_EEM_Model_Dimensions_PDM_Part8 ◊ Claim_EEM_Model_Exception_Event_Fact_PDM ◊ Claim_EEM_Model_Financial_Event_Fact_PDM ◊ Claim_EEM_Model_Header_Dental_Event_Fact_PDM ◊ Claim_EEM_Model_Header_Professional_Event_Fact_PDM ◊ Claim_EEM_Model_Lineitem_COB_Adjustment_PDM ◊ Claim_EEM_Model_MSQ_Event_Fact_PDM	**Note:** For legibility, the Dimensions data model for the Claim domain was divided into eight parts.

Related Schema	Related Data Model	Comment
	◊ Claim_EEM_Model_Override_Exception_Event_Fact_PDM ◊ Claim_EEM_Model_Previous_Location_Event_Fact_PDM ◊ Claim_EEM_Model_Provider_Event_Fact_PDM ◊ Claim_EEM_Model_Surgical_Procedure_Event_Fact_PDM • **Claim – DDI Release 7** ◊ Claim Financial Header Event_Logical ◊ Claim Financial Header Event_Physical ◊ Claim Override Explanation of Benefits Event_Logical ◊ Claim Override Explanation of Benefits Event_Physical ◊ Claim Pharmacy Detail Event_Logical ◊ Claim Pharmacy Detail Event_Physical ◊ Claim Pharmacy Header Event_Logical ◊ Claim Pharmacy Header Event_Physical ◊ Claim Professional Detail Event_Logical ◊ Claim Professional Detail Event_Physical ◊ Claim Professional Header Event_Logical ◊ Claim Professional Header Event_Physical ◊ Claim Provider Event_Logical ◊ Claim Provider Event_Physical ◊ Claim Related History Event_Logical ◊ Claim Related History Event_Physical ◊ Claim Warrant Event_Logical ◊ Claim Warrant Event_Physical ◊ Claim Condition Event_Logical ◊ Claim Condition Event_Physical ◊ Claim Coordination of Benefits Event_Logical ◊ Claim Coordination of Benefits Event_Physical ◊ Claim Dental Detail Event_Logical ◊ Claim Dental Detail Event_Physical ◊ Claim Dental Header Event_Logical ◊ Claim Dental Header Event_Physical ◊ Claim Diagnosis Event_Logical ◊ Claim Diagnosis Event_Physical ◊ Claim Exception Event_Logical ◊ Claim Exception Event_Physical ◊ Claim Facility Detail Event_Logical ◊ Claim Facility Detail Event_Physical ◊ Claim Facility Header Event_Logical ◊ Claim Facility Header Event_Physical	

Related Schema	Related Data Model	Comment
	◊ Claim Financial Detail Event_Logical ◊ Claim Financial Detail Event_Physical • Client ◊ Client_EEM_Address_Event_Fact_PDM ◊ Client_EEM_Authorized_Representative_Address_Event_Fact_PDM ◊ Client_EEM_Copay_Event_Fact_PDM ◊ Client_EEM_Eligibility_Event_Fact_PDM ◊ Client_EEM_Enrollment LockIn_Event_Fact_PDM ◊ Client_EEM_Enrollment_Event_Fact_PDM ◊ Client_EEM_Event_Fact_PDM ◊ Client_EEM_Level_of_Care_Event_Fact_PDM ◊ Client_EEM_Long_Term_Care_Liability_Event_Fact_PDM ◊ Client_EEM_Medicare_Part_C_Enrollment_Event_Fact_PDM ◊ Client_EEM_Medicare_Part_D_Span_Event_Fact_PDM ◊ Client_EEM_Merge_Event_Fact_PDM ◊ Client_EEM_Pregnancy_Event_Fact_PDM ◊ Client_EEM_Relationship_Head_of_Household_Event_Fact_PDM • **Client – DDI Release 7** ◊ Client Address Event - Logical ◊ Client Address Event - Physical ◊ Client Authorized Representative Address Event - Logical ◊ Client Authorized Representative Address Event - Physical ◊ Client Authorized Representative Event - Logical ◊ Client Authorized Representative Event - Physical ◊ Client Copay Event - Logical ◊ Client Copay Event - Physical ◊ Client Eligibility Event - Logical ◊ Client Eligibility Event - Physical ◊ Client Enrollment Event - Logical ◊ Client Enrollment Event - Physical ◊ Client Enrollment LockIn Event - Logical ◊ Client Enrollment LockIn Event - Physical ◊ Client Level Of Care Event - Logical ◊ Client Level Of Care Event - Physical ◊ Client Event- Logical ◊ Client Event - Physical	

Related Schema	Related Data Model	Comment
	◊ Client Merge Event - Logical ◊ Client Merge Event - Physical ◊ Client Pregnancy Event - Logical ◊ Client Pregnancy Event - Physical ◊ Client Relationship And Head Of Household Event - Logical ◊ Client Relationship And Head Of Household Event - Physical ◊ Long Term Care Patient Liability Event - Logical ◊ Long Term Care Patient Liability Event - Physical ◊ Medicare Enrollment Event - Logical ◊ Medicare Enrollment Event - Physical ◊ Medicare Part D Enrollment Event – Logical ◊ Medicare Part D Enrollment Event - Physical • **Drug Rebate – DDI Release 7** ◊ NMDS Drug Rebate Payment Check Log Event - Physical ◊ NMDS Drug Rebate Payment Event - Logical ◊ NMDS Drug Rebate Payment Event - Physical ◊ NMDS ROSI PQAS Event - Logical ◊ NMDS ROSI PQAS Event - Physical ◊ NMDS Drug Rebate A_R Event - Logical ◊ NMDS Drug Rebate A_R Event - Physical ◊ NMDS Drug Rebate Claim Event - Logical ◊ NMDS Drug Rebate Claim Event - Physical ◊ NMDS Drug Rebate Invoice Detail Event - Logical ◊ NMDS Drug Rebate Invoice Detail Event - Physical ◊ NMDS Drug Rebate Payment Check Log Event - Logical • Finance ◊ Finance EEM Model PDM • **Finance – DDI Release 7** ◊ Financial Payables Event - Logical ◊ Financial Payables and Receivables Event - Physical ◊ Financial Payables and Receivables Event - Logical ◊ Financial Header Event - Physical ◊ Financial Header Event - Logical ◊ Financial Claim Event - Physical ◊ Financial Claim Event - Logical ◊ Finance Receivables Event - Physical ◊ Finance Receivables Event - Logical ◊ Financial Payables Event - Physical • MCO	

Related Schema	Related Data Model	Comment
	◊ MCO_Allotment_Details_Event_Fact_PDM ◊ MCO_EEM_Cohort_Age_Event_Fact_PDM ◊ MCO_EEM_Cohort_Category_Of_Eligibility_Event_Fact_PDM ◊ MCO_EEM_Dimensions_PDM ◊ MCO_EEM_MCO_Detail_Event_Fact_PDM ◊ MCO_EEM_Plan_Detail_Event_Fact_PDM ◊ MCO_EEM_Plan_Rate_Event_Fact_PDM ◊ MCO_EEM_Plan_Specialty_Exclusion_Event_Fact_PDM • **MCO – DDI Release 7** ◊ Managed Care Plan Detail Fact - Physical ◊ Managed Care Plan Detail Fact -Logical ◊ Managed Care Plan Specialty Exclusion and Rate Event - Logical ◊ Managed Care Plan Specialty Exclusion and Rate Event - Physical ◊ MCO Detail Fact - Physical ◊ MCO Detail Fact- Logical ◊ Managed Care Allotment Event - Logical ◊ Managed Care Allotment Event-Physical ◊ Managed Care Cohort Category of Eligibility Event - Logical ◊ Managed Care Cohort Category of Eligibility Event - Physical • Prior Authorization ◊ Prior Authorization EEM Schema Detail ◊ Prior Authorization EEM Schema Header ◊ Prior Authorization Exception EEM Model • **Prior Authorization – DDI Release 7** ◊ Prior Authorization Header Event - Logical ◊ Prior Authorization Header Event - Physical ◊ Prior Authorization Notes Event - Logical ◊ Prior Authorization Notes Event - Physical ◊ Prior Authorization Exception Event - Logical ◊ Prior Authorization Exception Event - Physical • Provider ◊ Provider_EEM_Dimensions_PDM ◊ Provider_EEM_Provider_Address_Event_Fact_PDM ◊ Provider_EEM_Provider_Clia_Certification_Event_Fact_PDM	

Related Schema	Related Data Model	Comment
	◊ Provider_EEM_Provider_Disproportionate_Share_Event_Fact_PDM ◊ Provider_EEM_Provider_Electronic_Fund_Transfer_Event_Fact_PDM ◊ Provider_EEM_Provider_Enrollment_Event_Fact_PDM ◊ Provider_EEM_Provider_Event_Fact_PDM ◊ Provider_EEM_Provider_License_Certificate_Event_Fact_PDM ◊ Provider_EEM_Provider_Notes_Event_Fact_PDM ◊ Provider_EEM_Provider_Npi_Event_Fact_PDM ◊ Provider_EEM_Provider_Owner_Employee_Event_Fact_PDM ◊ Provider_EEM_Provider_Review_Claim_Event_Fact_PDM ◊ Provider_EEM_Provider_Review_Program_Event_Fact_PDM ◊ Provider_EEM_Provider_Taxonomy_Event_Fact_PDM • **Provider – DDI Release 7** ◊ Provider Event Fact - Logical ◊ Provider Event Fact - Physical ◊ Provider License Certificate Event Fact - Logical ◊ Provider License Certificate Event Fact - Physical ◊ Provider Notes Event Fact - Logical ◊ Provider Notes Event Fact - Physical ◊ Provider NPI Event Fact - Physical ◊ Provider NPI Event Fact - Logical ◊ Provider Owner Employee Event Fact - Logical ◊ Provider Owner Employee Event Fact - Physical ◊ Provider Review Claim Event Fact - Logical ◊ Provider Review Claim Event Fact - Physical ◊ Provider Review Program Event Fact - Logical ◊ Provider Review Program Event Fact - Physical ◊ Provider Taxonomy Event Fact - Logical ◊ Provider Taxonomy Event Fact - Physical ◊ Provider Address Event Fact - Logical ◊ Provider Address Event Fact - Physical ◊ Provider Clia Certification Event Fact - Logical ◊ Provider Clia Certification Event Fact - Physical ◊ Provider Disproportionate Share Event Fact - Logical ◊ Provider Disproportionate Share Event Fact - Physical	

Related Schema	Related Data Model	Comment
	◊ Provider Electronic Fund Transfer Event Fact - Logical ◊ Provider Electronic Fund Transfer Event Fact - Physical ◊ Provider Enrollment Event Fact - Logical ◊ Provider Enrollment Event Fact – Physical • **Statewide Human Resources Accounting Reporting Enterprise (SHARE) – DDI Release 7** ◊ SHARE GL Journal Fact - Logical ◊ SHARE GL Journal Fact - Physical • TPL ◊ Third Party Liability EEM PDM	
STG_FA_INPUT	• Flexible Analytics – Input ◊ FA_Admission_STG_FA_INPUT_PDM ◊ FA_Disease Staging_STG_FA_INPUT_PDM ◊ FA_Stage_INPUT_PDM	
STG_FA_OUTPUT	• Flexible Analytics – Output ◊ FA_Admission_STG_FA_OUTPUT_PDM ◊ FA_Disease Staging_STG_FA_OUTPUT_PDM	
FA_MART	• Flexible Analytics ◊ FA_Admission_MART_PDM ◊ FA_Disease Staging_MART_PDM	

Appendix A: ACRONYMS

The table below contains the acronyms used throughout this document as well as the definitions.

Table 11: Acronyms

Acronym	Definition
BI	Business Intelligence
CMS	Centers for Medicare & Medicaid Services
COTS	Commercial Off-the-Shelf
DBMS	Database Management System
DDI	Design, Development, Implementation
DED	Deliverable Expectations Document
DMS	Data Management Strategy
DS	Data Services
DW	Data Warehouse
EDA	Enterprise Data Architecture
EEM	Enterprise Event Model Schema
ERD	Entity Relationship Diagram
ETL	Extract, Transform, and Load
FA	Flexible Analytics
FK	Foreign Key
FHIM	Federal Health Information Model
FHIR	Federal Health Information Requirements
HHS	Health and Human Services
HSD	Human Services Department
IBM	International Business Machines Corporation
IGC	Information Governance Catalog
INT	Integrated Schema
KPI	Key Performance Indicator
LDM	Logical Data Model

Acronym	Definition
MARS-E	Minimum Acceptable Risk Standards for Exchange
MCO	Managed Care Organization
MECT	Medicaid Enterprise Certification Toolkit
MITA	Medicaid Information Technology Architecture
MMISR	Medicaid Management Information System Replacement
NIEM	National Information Exchange Model
O&M	Operations and Maintenance
ODS	Operational Data Store
PDF	Portable Document Format
PK	Primary Key
PMO	Project Management Office
PSA	Persistent Storage Area
QA	Quality Assurance
RDBMS	Relational Database Management System
REM	Reports and Extracts Schema
RFP	Request for Proposal
SFTP	Secure File Transfer Protocol
SHA-1	Secure Hash Algorithm 1
SHARE	Statewide Human Resources Accounting Reporting Enterprise
SI	System Integrator
SME	Subject Matter Expert
SMR	System Migration Repository
SQL	Structured Query Language
SRC	Source Schema
SSP	System Security Plan
STG	Staging Schema
TPL	Third-Party Liability

Acronym	Definition
XLC	eXpedited Lifecycle
XMI	XML Interchange
XML	eXtensible Markup Language

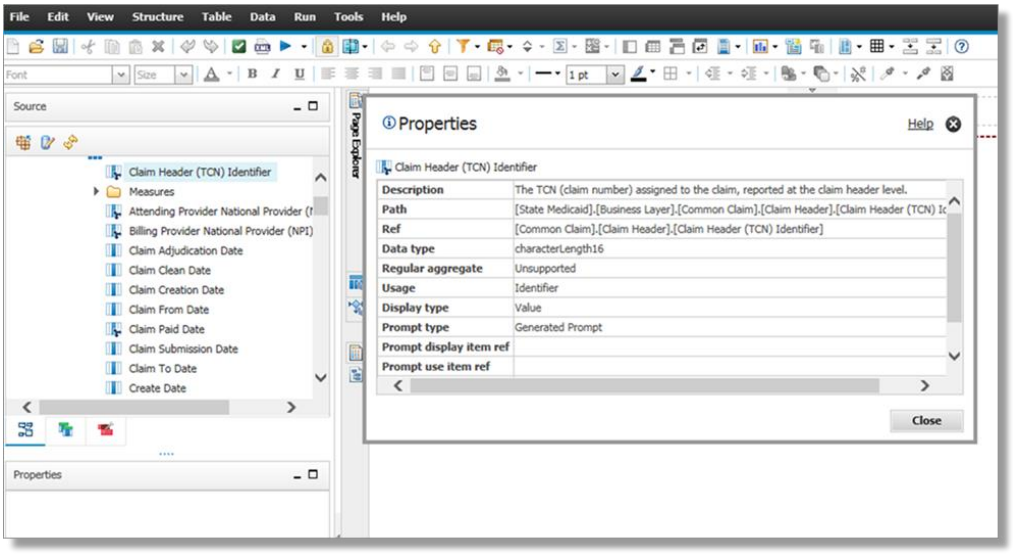
Appendix B: IBM PROPOSAL REQUIREMENTS

Table 12: IBM Proposal Requirements

RFP Category	RFP Description
Objective Category	System Design and Architecture
Requirement Number	3.29
Requirement	Offeror shall describe how its proposed Solution will optimize the retrieval and utilization of multi-dimensional data.
IBM Proposal Response	The NMDS solution optimizes the retrieval and utilization of multi-dimensional data. Conformed dimensions are the glue that holds together the entire data model in NMDS. They tie together data from disparate data sources into a comprehensive whole. They enable users to drill across related business entities and concepts. This drill-across analysis often yields the most interesting analyses. By designing our data models using thin, vertical slices rather than large, complex, horizontal layers (the modular design principles articulated by the MITA architecture), each slice provides end-user functionality and is fully testable on its own. In this way, data sets can be added, removed, or enhanced in a plug-and-play manner that does not disrupt the common data access layer. To ensure we are developing a good dimensional model, we identify: • **Facts:** We first identify the KPIs and numeric metrics to be analyzed and group them into fact tables; example includes metrics around claims. • **Dimensions:** We then define the primary reference data, defining the dimensions by which metrics are analyzed into dimension tables that are directly related to the fact table; examples include time, facility, location, client, and provider. We also must determine the logical hierarchies and groupings of the dimensions, such as “time dimension” extending from day to month to quarter to year. • **Aggregation:** We assume all data will be stored in the enterprise data store at its most granular level. We still must identify any specific aggregations that will be analyzed frequently. Typically, we begin with minimal aggregation and let the data access paths and performance determine the specific aggregates to be defined. This information helps with the modeling for the analytic data store and data marts.

RFP Category	RFP Description
	• **Historical Data:** It is important we understand the historical requirements and how changes should be handled for business analysis for each dimension; this is so we can determine how to handle changes to dimensions—Type 1, Type 2, or Type 3 SCD (Slowly Changing Dimensions). We also determine the audit-trail requirements. Following the above best practices, we will deliver a solution that empowers your staff to access the multi-dimensional data and create meaningful reports, quickly and easily.
RFP Category	**RFP Description**
Objective Category	System Design and Architecture
Requirement Number	3.42
Requirement	Offeror shall describe how its proposed Solution will import and maintain State-identified reference data bases (e.g., Census or Medicare data, commercial insurance data sets, etc.).
IBM Proposal Response	Reference data is an often overlooked but a vital part of a well-functioning enterprise data warehouse system. Truven Health manages reference data to help analysts produce accurate, credible information. Standard practice in healthcare BI is to store reference data from the source system. We follow that standard practice. We will store the relevant reference data for current and future data sources in the NMDS data warehouse. For example, we will include reference data used by the MMIS for claim processing, such as codes in use (procedure, diagnosis, drug, and others), fee schedules, and reimbursement data (OPPS, DRGs, RUGs, other fee-for-service payment limits). We store the reference data for point-in-time reference and maintain changes to the reference data whenever the MMIS is changed. We have cross-market experience using CMS data sources such as the National Plan and Provider Enumeration System; Medicare Parts A, B, and D data; and Healthcare Provider Cost Reporting Information System (HCRIS). Also, our health services researchers routinely use data sets from the U.S. Census Bureau (including American Community Survey and Current Population Survey), Behavioral Risk Factor Surveillance System (BRFSS), Pregnancy Risk Assessment Monitoring System (PRAMS), Corrections Data, Hospital Discharge Data, and many others to profile health services demand, supply, expenditures, quality, and outcomes. An example is the recently completed study of the Rhode Island Behavioral Health system.

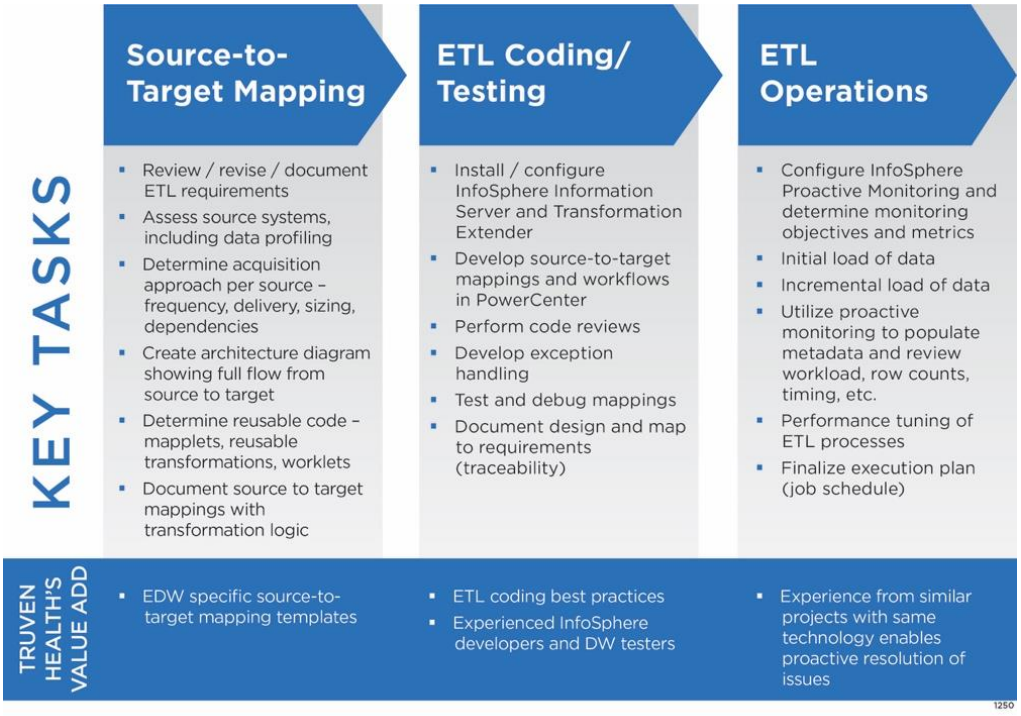
RFP Category	RFP Description
	To ensure data integrity, the reference tables are protected against being overwritten. Only the reference data maintenance processes will have write access to this data. All other users, processes, and systems will have read access to the data, as it is established in the DBMS security subsystem. This ensures the quality of the data is maintained and ready for use whenever it is needed. The data warehouse makes the reference data available for analysis, not just look-up: for example, users can run quantitative analyses on how current upper payment limits compare to past payment rates.
RFP Category	**RFP Description**
Objective Category	System Design and Architecture
Requirement Number	3.45
Requirement	Offeror shall describe how its proposed Solution will include tools that establish and manage a logical data model, and will include standards, responsibilities, relationships, definitions, domains, keys and entity-relationship diagrams (ERDs).
IBM Proposal Response	To support the data management initiatives of NMDS and the Decision Support Services Team, a proven data model and trusted data warehouse is a necessity, integrating all Medicaid enterprise data into a single version of the truth through data modeling techniques that minimize duplication and resolve ambiguity in relationships. Truven Health follows a progression of development of conceptual, logical, and physical data models, and we use the industry-leading tool, CA ERwin Data Modeler, to develop and maintain them. ERwin enables visualization and manipulation of complex data structures, streamlining the design process and synchronizing the model with the database design. With Erwin, we implement data modeling standards; define entity relationships; include data element definitions; create domains and keys; and build entity-relationship diagrams (ERDs). In addition, ERwin provides relational support for Oracle objects and allows both forward and reverse engineering capabilities. This tool also provides development objects, such as file layouts and Data Definition Language (DDL) objects. We use Erwin to build the Logical Data Model, which is a representation of business concepts laid out in visual form (ERD) that clearly shows these concepts and their various relationships. It is independent of the underlying physical implementation. Features of the logical data model:

RFP Category	RFP Description
	• Each entity in the logical data model is assigned a primary key—the attribute or set of attributes that distinguish one instance of the entity from another. • All the attributes for each entity are included. Relationships between entities are represented through foreign keys associated with the primary keys of the referring entity. • Normalization decisions are finalized in the logical data model resulting in the final normalized representation of entity-to-entity relationships. • Designing a logical data model that fits the State's needs is crucial, not just because it reflects the commitment to treat data as a true enterprise asset, but also because it enables efficient and effective storage of data that businesses can readily access. Conversely, a poorly designed model negatively affects many data warehouse components, making any rework more difficult. The logical data model affects: • The design of data integration programs, data standards, and enterprise application integration, degree of normalization, use of surrogate keys, and cardinality • Reference data • Business and technical metadata and the metadata repository As an added benefit, we have developed automated processes that easily share the business metadata across different tools. For example, we can take the field and table definitions initially captured in the CA ERwin data modeling tool and propagate those to the various end user tools, as shown in Figure 31. Thus, business analysts and other report developers will have access to the data dictionary directly in the reporting tool without having to refer to an outside document. This facilitates report development and improves communication about what is available for reporting in the BI tool.  **Figure 31. Propagation**

RFP Category	RFP Description
	This illustrates how the information in ERwin is propagated to Cognos Truven Health has extensive experience in modeling the entity relationships of Medicaid enterprises and has developed an in-depth understanding of the data that drives and is needed to support the delivery of care, and the payment and monitoring of the associated healthcare claims. Our in-depth knowledge combined with the power of our selected industry-leading COTS data modeling tool ensure that we can deliver the data models needed to seamlessly integrate the HSD enterprise.
RFP Category	**RFP Description**
Objective Category	System Design and Architecture
Requirement Number	3.47
Requirement	Offeror shall describe how its proposed Solution will inform/extend the data models in the DS and SI Architecture, including metadata management, to minimize maintenance tasks when additional data elements are produced or required by the Enterprise.
IBM Proposal Response	The NMDS solution is inherently designed to support changes to the data model with minimal impact to the overall system and reduced maintenance effort as HSD's needs evolve to meet everchanging business requirements. For example, our solution supports rapid and reliable integration of new HSD programs through a modular data model design, solid Governance and SDLC processes, and effective Metadata Management. Additionally, our solution offers Business Intelligence features that allow users to securely extend the data model to meet their analytic needs without dependency on IT resources. The following concepts describe how the NMDS solution meets your requirements. **Flexible, Modular Data Model Design** NMDS allows integrated data sets to be added, updated, and removed in a way that does not disrupt the common data access layer. By designing our data models using thin, vertical slices rather than large, complex, horizontal layers (the modular design principles articulated by the MITA architecture), each vertical slice provides end-user functionality and is fully testable on its own. This means that when the State wants to implement a new program, new data sets can be added in a plug-and-play manner that does not disrupt the common data access layer. **Governance and SDLC** For the NMDS to be successful, it must grow and change to meet your evolving data needs. Once it becomes clear you have new needs such as, adding a new HSD program,

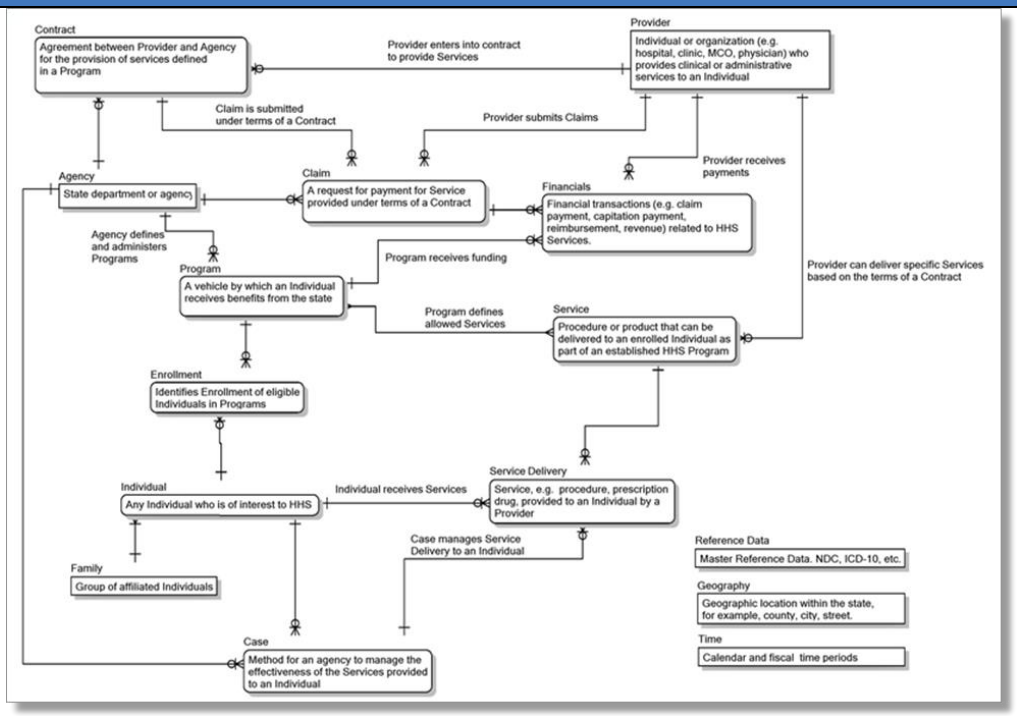
RFP Category	RFP Description
	the Data Governance Team prioritizes them based on their business value to the organization and perceived level of difficulty to implement them. From there, details of the requirement are fleshed out through the standard SDLC process. This includes identifying the system from which the data can be sourced and profiling the source system to determine the data's unique characteristics and quality. We coordinate with the SI and develop, code, and configure systems to operationalize the new requirement, and conduct extensive testing to ensure the new data is of sufficient quality to meet your organizational needs. Finally, to ensure adequate performance, we proactively monitor the production system as the new data enters it. Once in production, this new data often provides new business insights and raises questions that can be answered only by additional new data—and the cycle begins again. Through the application of governance and a proven SDLC, we ensure that new data elements are implemented correctly avoiding costly maintenance tasks that can arise otherwise. **Metadata Management** We understand that Data Governance and Metadata Management are extremely important topics to HSD. We utilize standards developed through our data modeling process to ensure that the proper metadata is captured and propagated downstream to the various analytic tools in the system. As such, stewards, analysts, and other stakeholders can quickly assess the impact of adding, updating, removing data to the system. For example, when integrating a new program, HSD can use metadata to determine if elements of the new data already exists in the system. Our process incorporates many descriptors from field type and size to description so that you will have complete understanding of what a particular field means and is used for. We deploy naming standards, data groupings and descriptions within your reporting solution which will provide an easy to use field identification process. Along with this, we enable within the reporting tool lineage functionality so that users can follow the complete lifecycle of data elements from source to report. The functionality that your new tool set contains will provide your users with the confidence that they are providing accurate reports. **Extending Through BI** Our solution, driven by Tableau and the sandbox, provides the ability to extend the NMDS data model with business intelligence function. Users will have the capability to import external data (excel, csv, etc.) as a data source. Once the data source is configured, numerical data can be aggregated using several predefined aggregations, such as Sum, Average, Mean, Count, Minimum, Maximum, Percentile, and more.

RFP Category	RFP Description
	Furthermore, this data can be joined with other data sources to bring more meaningful insights to the analysis and shared with others. In addition, the State can perform data-set additions and removals through use of the sandbox. The sandbox layer is a user-owned schema of the database that enables Truven Health staff and New Mexico super users to develop schemas that support their need to import external data and join it to any of the user-accessible data warehouse layers, given a linkable field exists. This enables super users to meet additional reporting needs, perform deep data analyses, complete highly sophisticated what-if reporting requirements, or prototype data from new HSD programs. Extending the data model through BI functionality offers HSD a means to quickly prototype, experiment with, and discover insights in new data sets without additional IT overhead and within little impact to the existing system.
RFP Category	**RFP Description**
Objective Category	System Design and Architecture
Requirement Number	3.48
Requirement	Offeror shall describe how its proposed Solution will have the capability to update or extend approved schemas/models to incorporate new data fields as needed or requested.
IBM Proposal Response	We have designed the NMDS solution with the flexibility to update and extend schemas and models to incorporate new data fields as needed. The development of your NMDS will be seeded from the Truven Health Core Medicaid Data Model. We built this model to industry standards, relying on our deep Medicaid and healthcare experience. The model itself is divided into subject areas like claims, provider, and client to both facilitate model management and make it easy to maintain. Using the Data vault modeling methodology of both Inmon and Kimball techniques, the model is split into a series of fact and dimension tables. These tables mirror the data relationships that exist in the business, and we have configured them to help you easily answer your business questions. Because of this modular approach to data model development, we can configure the model to meet the unique information needs of your organization, while at the same time remain compliant with industry standards. We can update and extend schemas, add new data sources, and fields, etc. to the model by simply mapping the new data to

RFP Category	RFP Description
	existing entities or, when necessary, creating new entities. Figure 32 illustrates the key tasks we follow as part of our ETL lifecycle.  **Figure 32. CMMI Level 3 Appraised Organization** *Our process addresses all aspects of the ETL lifecycle, from definition to implementation of new sources, fields, etc. and including ongoing maintenance of documentation.*

RFP Category	RFP Description
Objective Category	System Design and Architecture
Requirement Number	3.49
Requirement	Offeror shall describe how its proposed Solution will support physical-to-logical model mapping and rationalization of its data translation, data aggregation and data augmentation processes, and provide definition of model-to-model relationships of repository objects, data aggregation and flows utilizing graphical attribute-level mapping.

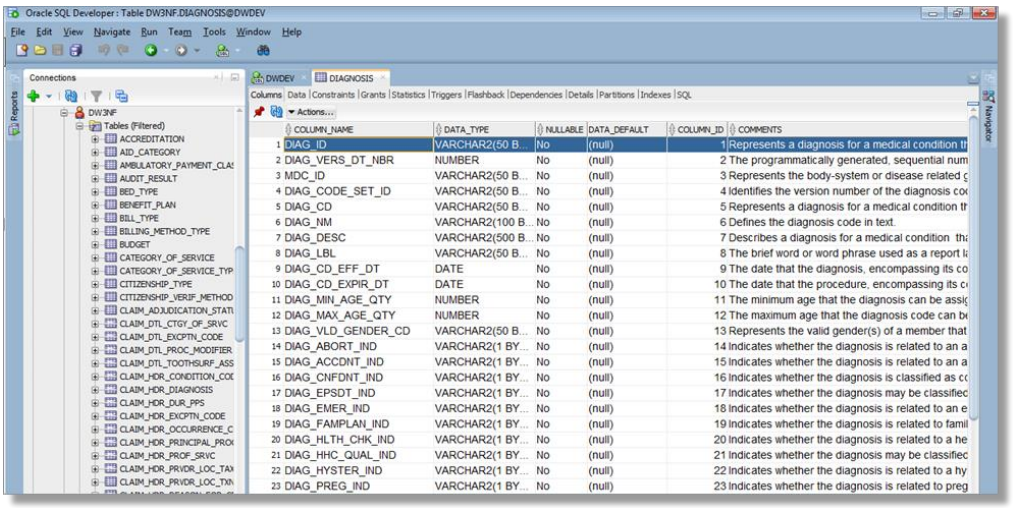
RFP Category	RFP Description
IBM Proposal Response	The NMDS solution will provide the necessary tools, documentation, and information to describe physical-to-logical model mapping, rationalization of data translations, data aggregation and augmentation processes, model-to-model relationships, and flows utilizing graphical and attribute level mapping. With our solution, New Mexico will be able to fully understand how your data is architected and organized in NMDS. The following sections walk through how we create and maintain data models, implement metadata tools, and leverage a Data Dictionary to meet your requirements. **Creation and Maintenance of Data Model** To support physical-to-logical model mapping and rationalization of its data translation, Truven Health follows a development progression of conceptual, logical, and physical data models. We use the industry-leading tool, CA ERwin Data Modeler, to develop and maintain them. ERwin enables visualization and manipulation of complex data structures, streamlining the design process and synchronizing the model with the database design. In addition, ERwin provides relational support for Oracle objects and allows both forward and reverse engineering capabilities. This tool also provides development objects, such as file layouts and Data Definition Language (DDL) objects. The **Conceptual Data Model** (Figure 33) identifies the highest-level relationships between different Medicaid entities and is developed consistent with the nature of Medicaid's business processes. The focus is on defining the major entities of interest in terms that are meaningful to the way the HSD conducts MITA business processes. A business process model defines a high-level relational map of the subject areas and cross-subject area dependencies for the data warehouse. The Conceptual Data Model captures these subject area relationships.

RFP Category	RFP Description
	 **Figure 33. The Truven Health Conceptual Data Model** *Allows users to easily see relationships between the various entities and how data flows between them.* The **Logical Data Model**, as described in 3.45, is a representation of business concepts laid out in visual form that clearly shows these concepts and their various relationships. It is independent of the underlying physical implementation. We then incorporate system and technical requirements, such as performance, to develop the physical design. Our goal is to maintain the logical model as much as possible. The logical model is transformed into a physical model that can be used to generate detailed Data Definition Language (DDL), including tablespaces, indexes, tables, comments, partitions, keys, views, and other Oracle objects. As changes are made to the model, they are made and versioned in ERwin. Once validated, the DDL changes are generated from ERwin and applied to the environment automatically or through scripts given to DBAs and system administrators in accordance with configuration management policies. The logical data model also serves as a foundation for data quality, as proper modeling of items such as domains and data types help validate data quality. Logical data models also must comply with data governance guidelines and enterprise data standards.

RFP Category	RFP Description
	The **Physical Data Model** represents how the model will be built in the database. A physical database model shows all table structures, including column name, column data type, column constraints, primary key, foreign key, and relationships between tables. Features of a physical data model include: • Specifications for all tables and columns • Foreign keys used to identify relationships between tables • De-normalization based on user requirements and performance considerations • Physical considerations causing the physical data model to be quite different from the logical data model • A physical data model that is different for different RDBMS **Metadata Management Tools** IT professionals such as business analysts, data analysts, and ETL developers use IBM InfoSphere Metadata Workbench (a component of the Information Governance Catalog) to explore and analyze relationships between information assets in the metadata repository. It is the central source from which to publish detailed knowledge about business and technical processes. InfoSphere Metadata Workbench provides IT professionals with a design-time tool for managing and understanding the assets that are generated and used by the IBM InfoSphere Information Server suite, and for extending that analysis to assets and processes that are external to the suite. By providing lineage reports, InfoSphere Metadata Workbench supports IT professionals who are responsible for compliance and governance initiatives that require lineage information (for example, Sarbanes Oxley or Basel II requirements). By providing an impact analysis that shows the effect of changes to information management environments, InfoSphere Metadata Workbench helps IT professionals to work most efficiently. InfoSphere Metadata Workbench supports these tasks: • Explore information assets in the metadata repository by using these features: • Details about the relationships of jobs, business intelligence (BI) reports, databases, data files, tables, columns, terms, stewards, servers, extended data sources, and other assets • Simple and advanced search and robust querying • Integrated cross-suite view of information assets • Details of extension mappings that describe external data flow and assets • Graphical view of asset relationships

RFP Category	RFP Description
	• Analyze dependencies and relationships of key assets and business BI reports by doing these tasks: • Trace lineage through jobs and databases to BI reports • Perform lineage analysis to understand where data comes from or goes to by using shared table information, job design information, operational metadata from job runs, and extension mappings • Perform impact analysis to understand dependencies and the effects of changes to a column or job in IBM InfoSphere Information Server and beyond • View operational metadata from job runs • Manage metadata to obtain in-depth analysis reports by doing these tasks: • Create and edit descriptions of information assets • Import or create assets that do not originate in InfoSphere Information Server: • Applications, stored procedures, and files that are defined as extended data sources • ETL processes that are defined as extension mappings • Assign terms, stewards, labels, and notes to information assets IBM® InfoSphere Metadata Workbench provides a visual, Web-based exploration of metadata assets that is generated, used, and imported by IBM InfoSphere Information Server. Access to these assets is controlled by built-in security functions so users can only see that information to which they have been granted access. InfoSphere Information Server components store design time, runtime, and glossary metadata in the metadata repository. Users can also import database and data file information into the metadata repository and create extended data sources and extension mappings that represent objects and processes that exist outside of InfoSphere Information Server. InfoSphere Metadata Workbench helps business and IT users explore and manage those metadata assets. The metadata workbench gives you reports on data flow, data lineage, and the impact of changes to data assets or physical assets. • **Business Metadata Repository**: InfoSphere tools provide a repository for all business glossary definitions; the interface allows multiple parties to collaborate and communicate about changes. Changes are not applied to the production repository until they have been reviewed and approved. Automatic communication can be delivered to interested parties. • **Technical Metadata Repository**: InfoSphere has a complete view of all technical metadata. When technical staff need to understand data structures, data relationships, or ETL processes, this repository gives them access to all necessary details.

RFP Category	RFP Description
	• **Quality Metadata Repository**: The metadata repository includes data collected by the InfoSphere QualityStage toolset. When combined with the operational metadata, it gives users a complete view into the accuracy, completeness, and timeliness of all data. Data Quality metadata will be captured during the life of the solution. The bulk of Data Quality metadata will be captured during the data acquisition and data integration processes performed against the incoming raw data. Most of this information will be generated by and accessible through the InfoSphere suite of data integration tools. • **Operational Metadata Repository**: The operational metadata will combine data from InfoSphere DataStage and the Metadata Management tool. All current processes and workflows can be viewed in detail in this repository. Access to the InfoSphere DataStage environment allows required schedule changes to be made easily. • **Security Metadata Repository**: Security metadata will be retrieved from a combination of sources; LDAP, DBMS, and reporting tools will have security components. The repository will be designed and built based on specific security metadata requirements. • **End User Metadata Repository**: This repository is not a standard component of the Metadata Management tool. However, all data required to provide this type of information will be available in the InfoSphere environment. As requirements are identified, the necessary end user metadata reports and dashboards can easily be built. • **Metadata Security:** All metadata in the system and databases will be secured as part of our standard security policies and procedures. Metadata within the database will be secured via authentication, role-based access and control, and object and row level security methods. Users will be provided access only to data. **Data Dictionary** Throughout the term of the contract, we will provide your users with an up-to-date, electronic, searchable, downloadable data element dictionary (DED) via our SharePoint environment. The DED gives information on the source, value, and characteristic attributes for data warehouse and data mart tables and data elements. We maintain the DED as a living document generated by the CA ERwin data modeling tool. Whenever changes occur to the model (either as part of a DDI release or a change in operations), we generate and upload the latest DED to the SharePoint site. If advanced users decide they want access to the actual physical or logical model definition, we provide that as well.

RFP Category	RFP Description
	In addition to the HSD SharePoint site, analytic sandbox and BI users will be able to access the DED directly within the tools they are using rather than having to refer back to it. We accomplish this by propagating the DED to end user tools, such as Oracle SQL Developer, Cognos, and Tableau, to ensure consistent understanding of the data elements regardless of the tool or environment the user is working in, as shown in the following DED Propagation exhibit. For example, ERwin provides relational support for Oracle objects and allows forward engineering capabilities to generate the Data Definition Language (DDL) objects used to create the physical model (Figure 34). Thus, your analytic sandbox users will be empowered to explore the data with a complete understanding to reveal new insights.  **Figure 34. DED Propagation** *This example shows how information in ERwin is propagated to Oracle.* Using the tools and functions that our solution offers, NMDS business users can get answers to questions such as "What are the sources for this report that I'm looking at?" and "Which transformations were applied to the data on which this report is based?" The answers to such questions are critical to the compliance laws currently in place in the healthcare industry. Similarly, IT analysts can get answers to questions regarding impact of change like "Which assets and processes would be affected if I change the constraints on this column?" These functions will deliver full transparency to HSD enterprise data, foster greater trust in HSD data, and help alleviate costs associated with not having the tools to understand your data fully.
RFP Category	**RFP Description**
Objective Category	System Design and Architecture

RFP Category	RFP Description
Requirement Number	3.51
Requirement	Offeror shall describe how its proposed Solution will minimize maintenance tasks when additional data elements are produced or required by the Enterprise.
IBM Proposal Response	The NMDS solution will minimize maintenance tasks when additional data elements are produced or required. Our solution architecture, including the proposed Truven Health Core Medicaid Data Model, is based on the CMS MITA Framework and guiding principles of the Seven Standards and Conditions. By designing our data models using thin, vertical slices rather than large complex horizontal slices, each vertical slice provides end user functionality and is fully testable on its own. In this way, data elements can be added, removed, or enhanced in a plug-and-play manner that does not disrupt the common data access layer. We can also add new data sources to the model by simply mapping the new data to existing entities or, when necessary, creating new entities, with little impact on the system and requiring minimal maintenance effort.
RFP Category	**RFP Description**
Objective Category	System Design and Architecture
Requirement Number	9.03
Requirement	Offeror shall describe how its proposed Solution will update report data fields as needed or requested and will provide for New Mexico-specific and user-defined report data fields for Enterprise use which will be revisable and expandable over time as necessary to support Enterprise program requirements.
IBM Proposal Response	In our proposed NMDS solution, authorized users will be able to update, at their discretion, report data fields in reports to which they have edit rights. To do so, a user simply logs into the system, navigates to the report that they wish to edit, opens it in the report authoring tool, and edits the field. For example, the report author could swap an existing attribute with a new attribute simply by dragging and dropping the new attribute over the existing attribute. Report authors will be able to create user-defined report data fields and calculations through our reporting solution's expression editor. If the user-defined field becomes widely used, the expression can be added to the NMDS sematic layer and published to authorized users in the HSD enterprise. We understand that the underlying data elements in the data warehouse will evolve over time. These changes will be carefully monitored, assessed, and implemented

RFP Category	RFP Description
	through a Data Governance process which we will help the State establish. The reporting and analytic tools in our solution help users prepare for and quickly adapt to changes in underlying data elements. For example, when a data change is identified, Cognos has the means to determine which reports reference the field identified for change so that you can determine the impact of making the data change. Once the impacted reports are identified, you can use our solution's usage reports to identify the users running the impacted reports and the frequency at which those reports are run to help prioritize effort, communicate changes, and minimize impact to downstream users. For Tableau users, our solution automatically manages metadata to adapt to a number of changes so that end users can continue to consume analysis without disruption. For example, changes to a published master metadata model can be automatically propagated to all the workbooks referencing the master model. Because of features like these, NMDS can maintain business intelligence independent of the source data to minimize disruption and provide greater flexibility.