



# **New Mexico Interagency Data Governance Council**

## **Data Quality Plan**

**Deliverable Owner: Frances Peralta-Donio**

**Data Governance Manager**

**Version 1.1**

**Date: December 12, 2022**

## Table of Contents

1.0	Introduction .....	4
2.0	Data Quality Plan Purpose .....	4
3.0	Data Quality Plan Scope .....	4
4.0	Approach .....	5
5.0	Culture and Governance .....	6
6.0	Process and Practices .....	8
6.1	Data Collaboration Team .....	8
6.2	Data Quality Standards .....	8
6.3	Metadata .....	10
6.4	Defects per Million .....	11
6.4.1	Source Data .....	12
6.4.2	Metadata .....	12
6.4.3	SI File Layouts and Process .....	13
6.4.4	DS Process .....	13
6.4.5	Quality Reporting .....	13
6.5	Criticality Level and Action .....	14
7.0	Roles and Responsibilities .....	16
8.0	Assumptions / Constraints / Risks .....	19
8.1	Assumptions .....	19
8.2	Constraints .....	19
8.3	Risks .....	19
9.0	Appendices .....	19
9.1	Appendix A: Deliverable Record of Changes .....	19
9.2	Appendix B: List of Acronyms .....	19
9.3	Appendix C: Terms and Definitions .....	20

## Table of Tables

Table 1 - Data Quality Plan Roles and Responsibilities .....	16
Table 2 - Deliverable Reference of Changes .....	19
Table 3 - List of Acronyms .....	19
Table 4 - List of Terms and Definitions .....	20

## Table of Figures

Figure 1: Multifaceted Approach for Data Governance ..... 5

Figure 2: Critical Components of a Data Quality Operating Model ..... 5

Figure 3: DGC Scope of Activities ..... 6

Figure 4: Data Governance Key Success Factors ..... 7

Figure 5: Data Governance Decision Matrix ..... 7

Figure 6: Data Collaboration Team RACI ..... 8

Figure 7: Data Asset Categories Prioritization ..... 9

Figure 8: Data Asset Categories Quality Objectives ..... 10

Figure 9: Metadata RACI ..... 11

Figure 10: DPMO Formula ..... 12

Figure 11: DPMO to Sigma Level Translation ..... 12

## 1.0 Introduction

New Mexico's (NM) Human Services Department (HSD) has adopted the Health and Human Services (HHS) 2020 vision, a transformational, enterprise-wide approach to the HHS business. NM HHS 2020 will move service delivery from a program-centric approach to a stakeholder-centric approach. NM 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 Data Quality Plan (DQP) defines the scope of data quality focusing on business outcomes, approach to evaluating, maintaining, and improving data quality, and the roles and responsibilities of all of us on this multifaceted approach to data quality.

## 2.0 Data Quality Plan Purpose

As we move into a data driven organization care and trust of our data assets is imperative to better serve our customers improving customer satisfaction and outcomes. This Data Quality Plan will be a living document to identify the various approaches and responsibilities of the enterprise to improve and sustain quality of our data assets.

## 3.0 Data Quality Plan Scope

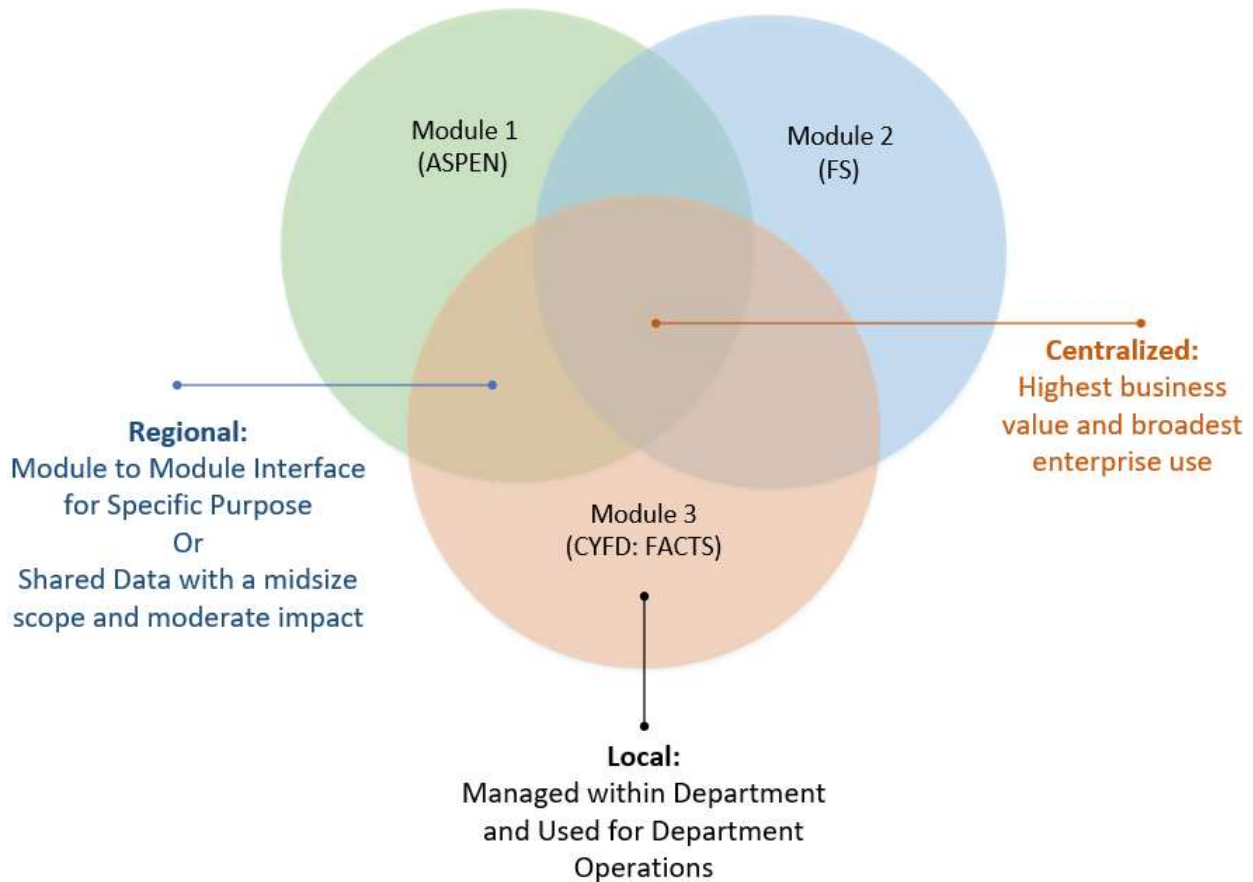
The scope for the Data Quality Plan is driven by business outcomes and use cases. The Departments strive to breakdown silos in their processes serving their collective customers to:

- Evaluate program outcomes based on value received
- Proactively target populations for 'at risk' and/or social vulnerability to improve security and promote independence
- Encourage change management to support value programs versus ongoing support of programs with no differential outcome
- Coordinate services between departments removing redundancy and improving cost effectiveness
- Perform internal processes at peak levels maintaining quality standards of service and acknowledgement by customers

The current use cases driving the approach the Data Quality Plan will begin with and mature include:

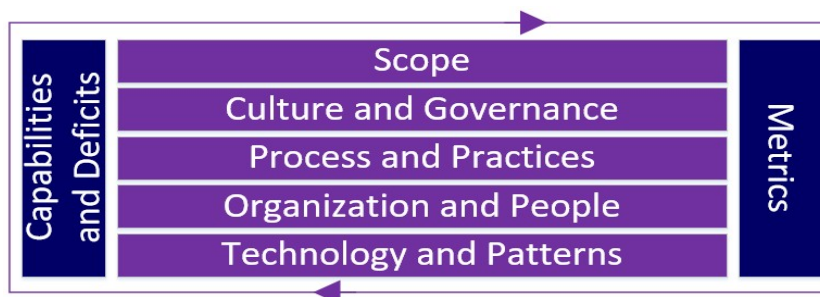
- Real-time module to module interfaces to support the coordination of services
- Data migration from various disparate systems into a single integrated data warehouse
- Leverage of integrated data to create an accurate customer record and 360° view of customer to improve security and promote independence will require:
  - Master data management
  - Metadata management
- Creation of algorithms to define impactful and evidence based programs
- Real-time operational analytics to improve staff effectiveness in serving customers

## 4.0 Approach



**Figure 1: Multifaceted Approach for Data Governance**

The Data Quality Plan takes a multifaceted approach in evaluating data quality and setting policies in how data shall be integrated to build trust in our data assets. Figure 1: Multifaceted Approach for Data Governance provides a visualization of the approach the Data Governance Council (DGC) will take depending on the situation. For example, the DGC defined the attributes used in identifying a unique person identifier and then the policies to create the golden record. The same decision would be different within a Department’s purview (aka local approach). Gartner, Inc. identifies several critical components for a Data Quality Operating Model (Figure 2: Critical Components of a Data Quality Operating Model) to help organizations build a comprehensive program in driving data quality assurance. As we grow in data quality maturity, we will incorporate additional components.



**Figure 2: Critical Components of a Data Quality Operating Model**

## 5.0 Culture and Governance

Data governance is a process to ensure the security and quality of data assets that facilitate a data driven organization. Data governance is both a defensive and offensive strategy. Defensive in reducing risk to protect the data in our care which contains confidential and valuable information. As an offensive strategy, it helps us become a data driven organization to better serve customers and creating evidence based programs leading to improved customer satisfaction, and when staff spends less time looking for, managing, and cleaning data we reduce redundancy and duplication of effort improving employee satisfaction.

The Multi-Agency DGC has representatives from HSD, Children’s Youth and Families Department (CYFD), Department of Health (DOH), Aging and Long-Term Services Department (ALTSD), and Early Childhood Education and Care Department (ECECD) with the responsibility of ensuring the security and quality of data assets. The DGC includes the following roles:

- Data Owners – Individuals from the business that make data asset decisions and enforce these decisions throughout the organization. They are ultimately accountable for the quality of the data as an asset, and identifying and classifying the security access levels.
- Data Custodians – Information technology individuals responsible for the technical environment and database structure to ensure safe custody, transport, storage, and implementation of business rules.

**Error! Reference source not found.**, depicts the scope of activities for the DGC.



**Figure 3: DGC Scope of Activities**

Working with the DGC are two other roles:

- Data Stewards – Individuals from the business that have thorough knowledge of their respective business area’s processes and procedures. These individuals ensure the data policies and data standards are adhered to in daily business and evaluate the data quality on a regular basis. They are subject matter experts for the data and its attributes.
- Business Intelligence Analysts – These are individuals who work with the data daily providing analytical expertise and accurate analytical solutions for business decisions. They maintain the

analytical foundation through business definitions and design of data marts to ensure integrity and reliability of analytics.

The HSD Data Governance Manager is the Data Governance Chair for the DGC which meets monthly. Currently, the HHS2020 Executive Steering Committee (ESC) is its governing body.

Gartner has identified nine key success factors for data governance (Figure 4: Data Governance Key Success Factors).



**Figure 4: Data Governance Key Success Factors**

The DGC has adopted the use of these key success factors as it works towards improving data quality, ensuring data security, and developing supporting policies and procedures.

The scope of the DGC is prioritized based on the enterprise use and data type characteristics including business value and volume to ensure adequate governance without adding burden. Figure 5: Data Governance Decision Matrix identifies the empowerment of the Data Owner(s) or Data Steward(s) within a single department or across divisions, or when decisions require the DGC review and approval.

	Business Measures / Dashboards	System Integrator, Master Data and Reference Data	Access to Data And Identifiable Data	Data Quality Review and Initiatives	
Single Division / Office					Decision remains with Data Owner and Data Steward
> 1 Division / Office in same Department					Along with DGC Chair, Data Owner and Data Steward of involved Divisions/Offices must review and approval
> 1 Department					Requires DGC review and approval
External					

**Figure 5: Data Governance Decision Matrix**

## 6.0 Process and Practices

Data systems, data flow, and data integration are complex systems and processes. Data quality initiatives is not a one and done deal as it must be comprehensive and agile without incorporating cost with a negative value. The processes and practices identified below will incorporate tasks such as data validation, data profiling, and quality monitoring.

### 6.1 Data Collaboration Team

Under the oversight of the DGC, a Data Collaboration team co-facilitated between the Data Governance Manager and the Database Administrator Lead is responsible for the process to extract data and integrate data to the enterprise data warehouse. The Data Governance Manager has overall accountability for the success of this team and shares responsibility with the Database Administrator Lead in managing daily activities of the Data Collaboration team. Figure 6: Data Collaboration Team RACI identifies the roles and responsibilities to effectively manage data / process defects.

Responsibilities	Data Governance Manager	Database Administrator Lead	Source System Data Owner	Source System Data Steward	BI Analyst	Source Module Data SME	SI Representative	Target Module Vendor (e.g. DS)
Inform DGC of major and critical defects including cause and planned solution. Seek DGC approval if necessary.	C,I	C,I	A	R	C,I	C,I	C,I	C,I
Target module identifies data issues. Target module will provide detail of data issue, including but not limited to, DPMO, and description of potential impact to their module. Information is entered into JIRA without PHI/PII/FTI.	C,I	C,I	I	C,I	I	C,I	I	A,R
Review, validation, and solution of data issue identified by target source. All information entered into JIRA ticket.	C,I	C,I	C,I	A,R	I	R	I	I
Approval / disapproval of solution in JIRA	R	C,I	A	C,I	I	C,I	I	I
Facilitate solution plan and process if necessary:								
Solution requires change in source system or source system data	C,I	C,I	A	R	C,I	C,I	I	I
Solution requires change in SI process	C,I	A	I	C,I	I	C,I	R	I
Solution requires change in Target module	C,I	A	I	C,I	C,I	I	I	R

Figure 6: Data Collaboration Team RACI

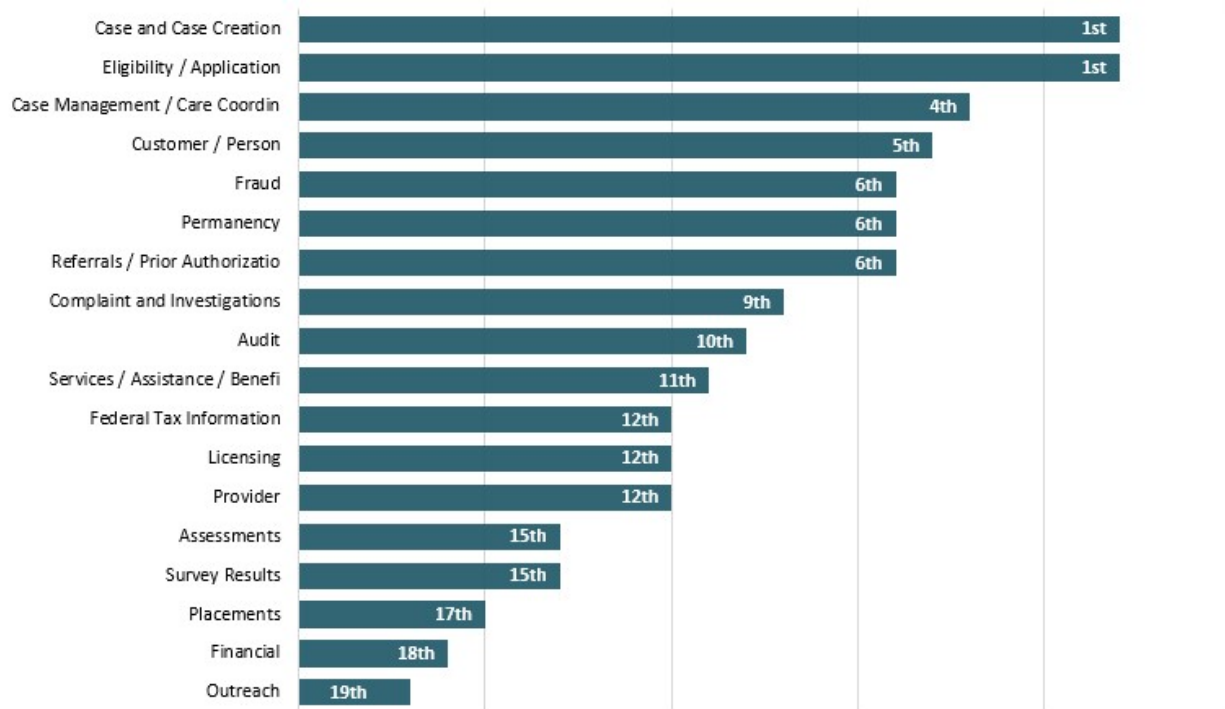
### 6.2 Data Quality Standards

Taking a proactive approach and evaluating data from source systems is time and cost effective to ensuring high data quality and preventing the propagation of a negative impact. The data quality standards provide direction to the data stewards on the what and how to measure data quality at the source system level. The data assets were for the collective agencies are placed into the following categories:

- Assessments
- Audit
- Case and Case Creation
- Case Management / Care Coordination
- Complaint and Investigations
- Customer / Person
- Eligibility / Application
- Federal Tax Information
- Financial
- Fraud
- Licensing
- Outreach
- Permanency
- Placements
- Provider
- Referrals / Prior Authorizations
- Services / Assistance / Benefits Administered
- Survey Results



The DGC evaluated each data asset category and prioritized them to identify which required the highest level of quality rigor because they bring the highest business value and broadest enterprise use. Figure 7: Data Asset Categories Prioritization presents the prioritization ranking of the data asset categories as identified by the DGC. This information tells the Data Stewards what data asset to measure first.



**Figure 7: Data Asset Categories Prioritization**

The DGC then evaluated the top five (5) data asset categories and identified the quality objectives the Data Steward will apply in descending order (the how). The quality objectives used are the following:

- **Accessibility** - Accessibility measurements ask questions about the estimated number of processes or personnel that can benefit from the data, and that are able to retrieve it.
- **Accuracy** - Accuracy measurements consider the degree to which the data values accurately describes real-world events or objects. This does not include missing data.
- **Completeness** - Completeness measurements considers the percentage of instances of data you have recorded versus the total available, or the percentage of missing fields in a record.
- **Integrity** - Integrity measurements consider the existence and correctness of linkages to and from a record, and often the legitimacy of relationships among attribute values within a record.
- **Timeliness** - Timeliness measurements consider the probability that a data record reflects its original source at any given time.
- **Uniqueness** - Uniqueness measurements consider the percentage of alternate or duplicative forms of data instance that exist.
- **Validity** - Validity measurements consider how well non-missing data accurately represents reality.

Figure 8: Data Asset Categories Quality Objectives provides the quality objective priorities for the top 5 data asset categories.

Cases and Case Creation	Eligibility / Applications	Case Management / Care Coordination	Customer / Person	Fraud
Validity	Uniqueness	Accessibility	Uniqueness	Accuracy
Accessibility	Accessibility	Accuracy	Accuracy	Integrity
Accuracy	Accuracy	Validity	Integrity	Validity
Integrity	Validity	Uniqueness	Completeness	Completeness
Completeness	Completeness	Completeness	Validity	Timeliness
Timeliness	Integrity	Integrity	Accessibility	Uniqueness
Uniqueness	Timeliness	Timeliness	Timeliness	Accessibility

Figure 8: Data Asset Categories Quality Objectives

### 6.3 Metadata

Metadata describes the content, structure, relationship, and characteristics of data. It brings business value in:

- classifying sensitive information and use that classification in setting access restrictions
- provides insight into what data is stored and combined with other data
- provides linkage back to the source and forward to mapped business intelligence tools, views, and data marts
- provides a way to easily identify downhill impact if any part of the data changes allowing business to proactively prevent negative impact due to changes

An enterprise modeling and data intelligence software tool will be leveraged to store the meta data catalog including the following information:

- Table name
- Field name
- Data type
- Data length
- Required Identifier / Nullable Flag
- Scale for numeric fields
- Precision for numeric field
- Description
- Description
- Primary Key
- Foreign Key
- Security Classification
- Default Value

The same enterprise modeling and data intelligence software tool will also store the mapping between source to target (called Lineage) including file layouts created by any Data Custodian tasked with the Extract, Transform, and Load (ETL) process. Figure 9: Metadata RACI depicts the metadata Responsible, Accountable, Consulted and Informed (RACI):

Responsibilities	Erwin Administrator	Data Steward	Data Owner	BI Analyst	Security Officer	Data Governance Manager	DB Administrator	Source System SME
Admin: Maintain Erwin site and application	R	I				I	A,R	
Admin: Identify user roles and related permissions in Erwin	R	C,I			A	C,I	C,I	
Admin: Identify users and assign appropriate roles	R	C,I			A	C,I	C,I	
Admin: Erwin administration to update/ manage users (resource manager)	A,R	I	I	I	C,I	C,I	C,I	I
Metadata: scan data sources / obtaining meta data from sources		R				A	C,I	C,I
Metadata: Organization of meta data files		A,R				C,I	C,I	I
Metadata: Maintain Metadata information. This includes table and field names, data type, length, precision, default value, nullable flag, security classifications, and primary and foreign keys		R			C,I	A	C,I	C,I
Metadata: Identification of security classifications		C,I			A,R	C,I	C,I	R
Business Terms: Define and maintain business terms. Terms are not the same as calculated but can be used to distinguish between interface and integration or abbreviation lookup		R	A			C,I	I	C
Mapping: Define and maintain mappings to BI tools		C,I		C,I		I	A,R	I
Mapping: Define and maintain mappings to SI ETL		R				A	R,C,I	C,I
Mapping: Define and maintain mappings to DW datamarts		R				A	R,C,I	C,I
Business rules: Define and maintain business rules		R	C,I	C,I		I	I	A,R
Business policies: Define and maintain business policies		R	A	C,I	R	I	I	R

Figure 9: Metadata RACI

## 6.4 Defects per Million

Data goes through several systems/processes and transitions before landing in the Enterprise Data Warehouse (EDW). To measure the process performance of the data migration, the DGC has selected a Six Sigma measurement to determine acceptable levels of performance. Six Sigma is a rigorous, focused and highly effective implementation of proven quality principles and techniques. (Pyzdek, 2010) 3.4 problems per million opportunities is a Six Sigma standard and response to the increasing expectations of customers and the increased complexity of modern products and processes. (Pyzdek, 2010) Defects per Million Opportunities (DPMO) is critical to quality measure which compares process observations against process requirements. Figure 10: DPMO Formula presents the DPMO formula which measures the ratio of the number of defects in one million opportunities, or how many times did you have a defect (flaw or mistake) for every opportunity there was to have a defect. Identified defects are filed and stored in JIRA where a team with knowledge of the source data, data governance, and module vendors collaborate on identifying both the cause and solution. The following sections detail the defect opportunities and section 6.5 specifies the criticality of the defect and resulting action.

$$\frac{\text{\# of Defects}}{\text{Defect Opportunities per Unit * Sample Size (\# of records in file)}} * 1,000,000$$

**Figure 10: DPMO Formula**

DPMO is used to measure quality of source data, the completeness and accuracy of the metadata, accuracy and completeness of SI file layouts, and the SI’s ETL process to transform data into a ready and acceptable state for Data Services (DS) consumption into the EDW. Figure 11: DPMO to Sigma Level Translation provides a translation from the DPMO into a Sigma level and corresponding accuracy level. The DGC evaluated the data areas and process which DPMO will be utilized as a quality indicator and determined the acceptable Sigma Levels for each. In determining the Sigma Levels, the DGC considered the impact to data quality, the age of data, and level of data stagnation. The Sigma Levels for each the data areas and processes are identified in the following sections.

### 6.4.1 Source Data

The SI will evaluate the source data based on business rules and metadata and provide defects to the data steward who will calculate the Sigma level. This will include the following opportunities for defects:

- Source system does not provide a unique index, primary key, or foreign key
- Duplicate records after primary keys are applied
- Mandatory fields with null values or empty spaces
- Source data feed does not match source data type
- Source data is missing file description
- Source data is missing field description

The DGC approved the following Sigma Levels for source data:

- Reference Files and Support Files = 5 Sigma
- Transactional Files (e.g. audit, transactional, service, eligibility)
  - >= 4 years old = 3 Sigma
  - 2 - 3 years old = 4 Sigma
  - 0 - 1 years old = 5 sigma

DPMO	Sigma Level	Accuracy
697,672	1	0.31
501,350	1.5	
308,770	2	0.69
158,687	2.5	
66,811	3	0.933
22,750	3.5	
6,210	4	0.994
1,350	4.5	
232.67	5	0.9997
31.69	5.5	
3.4	6	0.9999997

**Figure 11: DPMO to Sigma Level Translation**

In the event the source data has a sigma level below the DGC approved level, it is the responsibility of the Data Steward report to the DGC on findings, impact to analytics and operations, and remediation plan.

### 6.4.2 Metadata

The data steward will collaborate with the source system SME(s) and SI to validate the metadata from source system and SI layouts respectively. Together they will evaluate any of the following defects:

- File mapping is missing or incorrect

- Field mapping is missing or incorrect
- File description is missing or incorrect
- Field description is missing or incorrect
- Classification is missing or incorrect
- Classification does not have valid value
- Primary keys are missing or incorrect
- Foreign keys are missing or incorrect
- Data type is missing or incorrect
- Nullable indicator is missing or incorrect
- Scale is missing or incorrect for numeric fields
- Precision is missing or incorrect for numeric fields

The DGC approved a 5.5 Sigma (equates to 99.997% accuracy) for meta data. The Data Steward, in collaboration with the Source System Subject Matter Experts (SMEs), SI and DS will validate the metadata for any defects and provide regular reports to the DGC regarding such defects and resolutions.

### 6.4.3 SI File Layouts and Process

The SI file layouts contain information for the DS to extract the data for use in the EDW. The SI in collaboration with the data steward and DS will validate the file layout for potential defects covering:

- Layouts
  - Missing header names
  - Missing field(s)
  - Data type is missing or incorrect
  - Length is missing or incorrect
  - Nullable indicator is missing or incorrect
  - Scale is missing or incorrect for numeric fields
  - Precision is missing or incorrect for numeric fields
  - File acceptance criteria is missing
  - File does not reconcile to 'file acceptance criteria'
- Process
  - File is not consistent with layout
  - Data contains duplicate records as defined by primary keys
  - Field delimiters are not consistent or correctly applied
  - File header are not present or inaccurate
  - Mandatory fields have null values

### 6.4.4 DS Process

The DS vendor, will evaluate data received for the expected quality standards as defined in section above 'SI File Layouts and Process'. The DS may also establish additional data quality validations upon review and approval of the DGC. Any additional data quality validations must include impact to their module.

### 6.4.5 Quality Reporting

On a minimum monthly basis, each of the above areas identified in 6.3 will produce regular quality reporting identifying DPMO, origin of problems, potential impact to customers and processes as defined

in 6.4, and notification to Data Governance Manager and Data Steward of DPMO's. It is the responsibility of the Data Governance Manager, Data Owners and Data Stewards to review the quality reports on a regular basis and collaborate on a resolution and communication if necessary,

## 6.5 Criticality Level and Action

Data defects, whether created via data transmission or when created in the modules has a potential negative impact, of varying degrees, to ensuring customers receive appropriate benefits, operations receive information to make appropriate real-time decisions, and leadership has accurate analytics to made decisions impacting customers and staff. Four criticality levels are defined with subsequent actions to either monitor the defect, take immediate action in creating a work around, or changing processes or systems to remediate problems. The actions identified will occur if at least one bullet is true under the impact statement.

- **Minimal Impact**
  - Data defects for customers no longer covered by department programs.
  - Data defects are cosmetic in nature and do not affect data usage for either operations or analytics.
  - **Action:** Data Steward will evaluate impact of data defects and resources required to remediate. The Data Steward will take plan to appropriate DGC level for approval which may be to take no action.
- **Minor Impact**
  - Data Defects for current customers has no impact on receiving benefits.
  - Data defects are easily identified and temporary work around is available for analytics
  - There is no impact to operational reporting.
  - **Action:** Data Steward will collaborate with appropriate parties to develop a work around and collaborate with appropriate parities, including Data Owner and Data Custodian on a remediation plan. The Data Steward will take plan to appropriate DGC level for approval.
- **Major Impact**
  - Data defects for customers has potential to delay benefits for up to 5% of current or potential customers.
  - Data defects are identified but work around is complex in ensuring accurate analytics.
  - Operational reporting is negatively impacted and operations remains able to function in limited capacity or by leveraging other resources.
  - **Action:** Data Steward and Data Owner will lead collaboration to develop a plan in less than three (3) months and lead implementation to correct data defects. The plan is submitted to appropriate DGC level for approval and Data Owner will provide regular updates to full DGC. The Data Governance Manger will ensure ESC is informed and consulted if necessary.
- **Critical Impact**
  - Data defects for customers has potential delay to delay benefits for more than 5% of current or potential customers.
  - Data defects identified require changes in system or business processes to correct as impact to analytics is not overcome with temporary work arounds.
  - Data defects negatively impact operational reporting to the point they no longer support operations.
  - **Action:** Data Steward and Data Owner will lead collaboration to develop a plan in less than one (1) month and lead implementation to correct data defects. The plan is

submitted to appropriate DGC level for approval and Data Owner will provide regular updates to full DGC. The Data Governance Manger will ensure ESC is informed and consulted if necessary.

## 7.0 Roles and Responsibilities

Data quality is not just the responsibility of a Data Steward or the Data Governance Council, it is our collective responsibility to ensure the data is timely, accurate and complete. The following RACI focuses around several areas regarding data governance including business intelligence analytics, communication, DGC, process to create and enforce data policies, quality measurement, security, and tools. While the RACI does focus on the individuals directly linked to the DGC, it is important to note the Data Owner and Data Steward have direct responsibility over the security and quality of data in their direct control by the modules they use to run daily business operations.

**Table 1 - Data Quality Plan Roles and Responsibilities**

Category	Responsibilities	Data Owner	Data Steward	BI <sup>1</sup> Analyst	Data Custodian
BI Analytics	Analyze existing BI reporting processes and suggest improvements including developing protocols for optimal filtering and variance	I	C	A	R
BI Analytics	Assessing analytics on an ongoing basis and develop new data analysis processes as required	I	C	A,R	I
BI Analytics	Create policies and procedures regarding collating, combining, deriving, and analyzing data	I	C	A,R	I
BI Analytics	Define standard measures, algorithms, and data marts for use by other analysts to improve performance and facilitate repeatability and reproducibility.	I	C	A,R	C
BI Analytics	Prepare and maintain BI reporting and analysis documentations	I	I	A,R	I
BI Analytics	Provide recommendations to enhance BI system performance and capacity	I	I	C	A,R
Communication and Management	Communications and management of the following: – Applications’ functions and the data that is created, updated, or deleted within those applications – Impact of potential system or business process changes on the underlying data – Business rules for CRUD rights (Create, Read, Update, Delete) for data – How data is used operationally and/or analytically identifies triggers for change control	A	R	C,I	C,I
Communication	Communicates decisions to other stewards to ensure understanding and determine impacts.	C,I	A,R	C	I
Communication	Receives communications and directives from the DGC and passes them along to the respective members of their business groups	C,I	A,R	I	I



Data Quality Plan

Category	Responsibilities	Data Owner	Data Steward	BI <sup>1</sup> Analyst	Data Custodian
DGC	Escalates issues to the DG Facilitator that cannot be agreed upon and are expected to require decisions by the Data Governance Council	A,R	R	I	C,I
DGC	Participates in data governance sessions where proposed data-related changes are presented and analyzed.	A	R	C,I	R
DGC	Remove barriers and risks that impact the implementation and sustainability of the Data Governance program and conform to the conditions and expectations articulated by the DGC	A,R	C,I	C	R
Process	Comply with all Federal and State regulations policies applicable to the data in your custody	I	A	I	R
Process	Maintain disaster recovery plans and facilities appropriate and adequate meeting business needs to maintain or restart operations in the event systems or facilities are impaired, inaccessible, destroyed, or breached.	I	C	I	A,R
Process	Manage data user access as prescribed and authorized by Data Owner	C	C	I	A,R
Process	Responsible for implementation, maintenance and enforcement of data policies, standards, and compliance to governance processes	A	R	I	I
Process	Works closely with the DGC to evangelize Data Governance within business groups and attends DGC sessions and articulates the business problems that need attention. They are expected to contribute to the discovery process and help identify resources that would be needed to support the initiatives.	A,R	R	C,I	C,I
Quality	Defines processes and data reconciliation controls for maintaining the quality of data across the enterprise	A	R	I	C
Quality	Critically evaluate and screen data, including profiling to identify any issues or anomalies	A	R	C,I	C,I
Quality	Defines criteria for the integrity and quality of data definitions	A	R	C,I	C,I
Quality	Determines acceptable levels of data quality and monitors that quality on an ongoing basis by establishing data metrics based on key performance indicators (KPIs). This function will include those requirements affecting compliance	A	R	C	I
Quality	Ensures industry standards impacting data quality and definition are enforced for their particular domain	A	R	I	I
Quality	Helps identify specific measurements and validations that form the data standards	A	R	I	I

## Data Quality Plan

Category	Responsibilities	Data Owner	Data Steward	BI <sup>1</sup> Analyst	Data Custodian
Quality	Identifies events which should trigger Data Governance change control activities: – New fields are added to the application and supporting database – New screens, filters or views are added to the application – New online or offline methods for adding or maintaining data are added – New reports – New calculations – Business process or procedure changes – Regulatory changes impacting the organization	A	R	R	C,I
Quality	Implements process for data reconciliation and routine reporting of results	I	C	I	A,R
Quality	Manage data anomalies via data correction or implementation of tools and processes to prevent future occurrences.	A	R	C	C,I
Quality	Participates with SMEs and other team members in analyzing data-related issues and developing recommendations for resolving those issues	A	R	C	C,I
Quality	Supports the user community by collecting, collating, and triaging issues and problems related to data. Reports metrics and status of issues to the DGC as warranted	A	R	C	C,I
Quality	Validate the data warehouse system for data integrity, quality, and accuracy	I	C	I	A,R
Security	Enforces data security rules and works with the SMEs across Domains to identify and classify security access levels	A	R	I	C,I
Security	Identify and classify data security access levels. [Note: Requires consultation with security officer.]	A	R	I	I
Security	Identify who/what has access to data	A,R	C	I	I
Tools	Maintain physical and system security and safeguards appropriate to the classification level of the data in their custody	I	C	I	A,R
Tools	Manage installations, upgrades, modifications, maintenance, and troubleshooting of BI platforms and tools	I	I	C	A,R

<sup>1</sup> Business Intelligence Analyst

## 8.0 Assumptions / Constraints / Risks

### 8.1 Assumptions

### 8.2 Constraints

- The implementation of a data quality plan requires resources that may not be readily available or require additional work to current resources.
- Tools and/or access to data will prohibit the ability to successfully complete data quality measurements.

### 8.3 Risks

- Data quality measurement is a new concept and may reveal inadequate data which was used to support various strategic decisions.
- As data quality is a new concept to the project and organization, it will take time before HSD universally adopts the data quality plan, measurements, and learns to appropriately manage data defects.

## 9.0 Appendices

### 9.1 Appendix A: Deliverable Record of Changes

The deliverable will include a record of changes in the following form:

*Table 2 - Deliverable Reference of Changes*

Version Number	Date	Author/Owner	Description of Change
V1.0	9/22/2022	Frances Peralta-Donio	
V1.1	12/12/2022	Frances Peralta-Donio	Minor changes from DGC review Approval of Data Quality Plan by DGC

### 9.2 Appendix B: List of Acronyms

A list of project-specific acronyms will be maintained on the MMISR SharePoint site.

*Table 3 - List of Acronyms*

Acronym	Definition
ALTSD	Aging and Long Term Services Department
BI	Business Intelligence
CRUD rights	Create, Read, Update, Delete
CYFD	Children's Youth and Families Department
DGC	Data Governance Council
DOH	Department of Health

Acronym	Definition
DPMO	Defects Per Million Opportunities
DQP	Data Quality Plan
DS	Data Services
ECECD	Early Childhood Education Center Department
EDW	Enterprise Data Warehouse
ETL	Extract, Transform, and Load
HHS	Health and Human Services
HSD	Human Services Department
KPIs	Key Performance Indicators
NM	New Mexico
RACI	Responsible, Accountable, Consulted, and Informed
SI	System Integrator
SME	Subject Matter Experts

### 9.3 Appendix C: Terms and Definitions

*Table 4 - List of Terms and Definitions*

Term	Definition
Algorithm	A process or set of rules to be followed in calculations. For example, an algorithm can be created to identify clients without sufficient access to behavioral health care or identification of clients at high risk of diabetes.
Analytics – Business Intelligence	Business Intelligence analytics go beyond operations (what happened) and provide insight into why (diagnostic analytics), what is coming (predictive analytics), and suggestions what should be done or what can we do (prescriptive analytics)
Analytics - Operational	Analytics to support tactical operational functions, normally done in real-time and used to make instant decision in business operations to improve efficiency and streamline everyday operations.
Data Governance	Process to ensure the security and quality of data assets that facilitate a data driven organization.
Data Governance Role: Business Intelligence Analyst	Provide analytical expertise in providing accurate solutions for business decisions. Maintain the analytical foundation through business definitions and design of data marts to ensure integrity and reliability of analytics.
Data Governance Role: Custodian	Responsible for the technical environment and database structure to ensure safe custody, transport, storage, and implementation of business rules.
Data Governance Role: Data Owner	Individuals that make data asset decisions and enforce these decisions throughout the organization. Ultimately accountable for the quality of the data as an asset, and identifying and classifying the security access levels.
Data Governance Role: Data Steward	Individuals that have thorough knowledge of their respective business area's processes and procedures. Ensure the data policies and data standards are adhered to in daily business. SMEs for data entity and attributes.
Data Mart	Datamart, like a data warehouse is structured to support complex queries. The data in a data mart is a subset of the data warehouse to focus on a specific subject like inpatient stays, emergency room visits, customer demographics, and episodes.

Term	Definition
Data type	An attribute on a piece of data that tells a computer system how to interpret its value. Examples: integer, character, date, time
Data Warehouse and Enterprise Data Warehouse	A data warehouse is structured to support complex queries, integrate data from multiple sources, and enhance data to support a plethora of analytics.
Domain	Module or part of a module. E.g. Omnicaid (module), claims and providers can be part of a module
Foreign Key	A column or group of columns that provide a link between two different tables.
Master Data Management	Master data management is a business led and technology-enabled approach where the business and IT collaborate to ensure the uniformity, accuracy, stewardship, semantic consistency and accountability of the shared master data assets.
Metadata	Data about data assets that supports design for data integration, quality, lineage, privacy, security, and integrity.
Precision for Numeric Fields	Number of digits allowed in a number, for example 123.45 has a precision of 5.
Primary Key	Primary keys identify unique records in a data table.
Required Identifier / Nullable Flag	Identifies if a field is required and cannot be null. For example, for a client record the first name is required and cannot be blank or null.
Scale for Numeric Fields	Number of digits to the right of the decimal point.
Security Classification	Identifies the level of security applied to data which is used to provide or secure user access. Currently HSD has the following classifications: PHI, PII, PHI/PII, PII++ (Adoption Care), and Restricted.