National Data Index

Operational Excellence

Version 1.0

Document Control

Version	Revision Date	Contributor	Modification
1.0	Oct 2023	SDAIA	Final Version

Contents

1 Glossary

Term	Description
DMP	Data Marketplace
GSB	Government Service Bus
NDB	National Data Bank
NDL	National Data Lake
NDC	National Data Catalog
NDI	National Data Index
NDMO	National Data Management Office
NIC	National Information Center
ODP	Open Data Platform
RDP	Reference Data Management Platform
SDAIA	Saudi Data and Artificial Intelligence Authority

2 Introduction

The Saudi Data and Artificial Intelligence (SDAIA) is on the mission of developing a data-driven economy and fostering the data literacy level across the government entities for operational efficiency and decision-making in the Kingdom of Saudi Arabia (KSA). To this end, SDAIA has launched the National Data Index (NDI) framework that aims at measuring the efforts and progress in transforming the data into a vital economic resource for unlocking innovation, driving economic growth and transformation, and improving the national competitiveness in an organized and accelerated manner.

NDI encompasses the National Data Management Office (NDMO) Data Management and Personal Data Protection Framework prescribed domains [1] and is composed of three essential components: Compliance, Maturity, and Operational Excellence (OE). Each government entity will be measured on a regular basis across these three components.

This document introduces and elaborates the Operational Excellence (OE) component of NDI. OE measures the real advancement and achievement of efficiency and effectiveness in the government entities data management operation processes. It utilizes the required information captured from various national data platforms managed and operated by SDAIA (e.g., National Data Bank (NDB)[1]) to evaluate the efficiency and effectiveness of data management practices at the entities.

The OE component is built around measurable metrics that have been selected following various cycles of deliberations and are linked accordingly to the core objectives of NDI. This document describes the process and methodology of calculating

---

[1] The National Data Bank (NDB) is a constellation of objectively interconnected national data platforms (e.g., National Data Lake (NDL), Data Marketplace (DMP), National Data Catalog (NDC), and Open Data Platform (ODP)) that aim at accelerating the data sharing, improving the national data quality, and instilling data as a common denominator for the digital economy in the Kingdom.

the OE component, which is based on an empirical method focusing on identifying problems and improvement opportunities pertaining to the entities data management journey.

3 Operational Excellence Calculation Approach

The OE component consists of a number of metrics that are divided across the NDMO data management domains. Section 4.2 lists all the defined metrics and one such example is *DSI.OE.01* which is defined to measure the percentage of attributes shared by each entity on the Government Service Bus (GSB) that are not officially produced by the entity or delegated to it from the producer entity of those attributes. This metric is categorized under the Data Sharing and Interoperability (DSI) domain of NDMO.

In order to measure each metric, the national data platforms need to provide necessary metadata. For example, the required metadata from the Government Service Bus (GSB) and the National Data Catalog (NDC) will be acquired and subsequently used to calculate the metric *DSI.OE.01* as per its defined calculation method.

Since metric scores have different measurement units (e.g., time and percentage), a single numerical OE score is required for calculating the OE score of each entity. In order to normalize the measurement units, each metric output is mapped to a predefined scale that ranges in value from 0 to 5 where 0 signifies "Unacceptable" and 5 represents "Leader". The complete list of scale levels and values is as follows:

Scale Level	Scale Value
Unacceptable	0
Low	1
Fair	2

Scale Level	Scale Value
Good	3
Excellent	4
Leader	5

Once the metric scores are mapped to the unified scale, the OE score of an entity is calculated using the following weighted sum equation:

$$OE\ Score = \frac{\sum_{i=1}^{|L|} w_i * s_i}{|L|}$$

where $L = \{m_1, m_2, \dots\}$ is the list of metrics defined in Section 4.2, s_i represents the score of the entity for metric m_i, and w_i represents the weight of that metric. The rationale for introducing weights in the calculation is to administer the role of the metrics based on their importance, priority, and the dynamic nature of the data management landscape.

4 Operational Excellence Metrics

This section explains the structure of the OE metrics and the metrics that have been defined under each applicable domain of the NDMO Data Management and Personal Data Protection Framework.

4.1 Metric Structure

The metric structure consists of several elements discussed in the table below.

ID	Element Name	Description
1	Metric ID	A unique identifier of the metric using the following format [DomainID.OE.NUMBER] where NUMBER is an ordering number of the metric within the NDMO domain and OE. For

ID	Element Name	Description
		example, MCM.OE.01 is the first OE metric within the Data Catalog and Metadata domain of NDMO
2	Metric Name	The name of the metric; e.g., Response Time of GSB APIs
3	Metric Description	A high-level description with a detailed explanation and rationale for the metric
4	Domain Name	The name of the NDMO domain; e.g., Data Catalog and Metadata (MCM)
5	Data Platforms	The name of the national data platforms which will support the metric calculation
6	Definitions	The constituents of each metric required to calculate the output of the metric
7	Calculation	The process that transforms the metric's constituents into one output value using a predefined equation
8	Measurement Unit	A definite magnitude of the metric output
9	Acceptable Threshold	The acceptable value beyond which the metric output will not be acceptable
10	Scale Intervals	The intervals that correspond to the unified scale levels
11	Version History	The version history allowing versioning control of changes following the release of the document
12	Dependencies	The pre-requisites for calculating the metric output

4.2 List of Metrics

This section details the list of metrics of the OE component. Note that a number of NDMO domains are not included in the list due to the nature and significance of these domains and the unavailability of the quantitative metadata for the empirical observation and measurement of the related metrics.

4.2.1. Data Sharing and Interoperability (DSI)

Element Name	Element Details
Metric ID	DSI.OE.01
Metric Name	Attributes shared on GSB but not produced by the entity
Metric Description	This metric measures the percentage of attributes shared by the entity on GSB that are not officially produced by the entity or delegated to it from the producer entity of those attributes. In other words, this metric determines the proportion of attributes that should not be shared on GSB by the respective entity. This metric helps in achieving the required governance on data sharing and interoperability and ensures that only designated data producers or delegates can share the data on GSB, hence improving the quality of data in the Kingdom.
Domain Name	Data Sharing and Interoperability (DSI)
Data Platforms	Government Service Bus (GSB) and National Data Catalog (NDC)
Definitions	• Number of attributes shared on GSB that are not produced by the entity or delegated to it • Total number of attributes shared on GSB by the entity
Calculation	= Number of attributes shared on GSB that are not produced by the entity or delegated to it / Total number of attributes shared on GSB by the entity * 100
Measurement Unit	Percentage
Acceptable Threshold	10%
Scale Intervals	Unacceptable: > 10% Low: (8%, 10%] Fair: (6%, 8%] Good: (4%, 6%]

Element Name	Element Details
	Excellent: (2%, 4%]
	Leader: <= 2%
Version History	
Dependencies	

Element Name	Element Details
Metric ID	DSI.OE.02
Metric Name	Systems integrated with NDL
Metric Description	This metric measures the percentage of systems shared by the entity with the National Data Lake (NDL) against the total number of the entity's systems requested by the NDL team. Note that a system will be considered integrated only if all required dimensions of its data are fully sourced to NDL.
	This metric aims at accelerating the efforts in enriching NDL, which is the national data lake house of the Kingdom, with high-value and wide spectrum data assets generated by the various government entities. It also helps in achieving the goal of making NDL the unified single source of truth for analytical data assets.
Domain Name	Data Sharing and Interoperability (DSI)
Data Platforms	National Data Lake (NDL)
Definitions	• Number of systems integrated with NDL by the entity.
	• Total number of the entity's systems requested by NDL
Calculation	= Number of systems integrated with NDL by the entity / Total number of the entity's systems requested by NDL * 100
Measurement Unit	Percentage

Element Name	Element Details
Acceptable Threshold	70%
Scale Intervals	Unacceptable: <= 70% Low: (70%, 75%] Fair: (75%, 80%] Good: (80%, 85%] Excellent: (85%, 90%] Leader: > 90%
Version History	
Dependencies	

Element Name	Element Details
Metric ID	DSI.OE.03
Metric Name	Data sharing agreement processing
Metric Description	This metric measures the amount of time taken by the data producer to process the data sharing requests raised by the consumer entities. This metric considers the time taken for either approving or rejecting the requests as part of the processing time. This metric aims at improving the data sharing and accessibility by optimizing the overall data sharing agreement approval lifecycle. It also helps in accelerating the data consumption by lowering the barriers for accessing the data required for decision-making and insight generation.
Domain Name	Data Sharing and Interoperability (DSI)
Data Platforms	Data Marketplace (DMP)
Definitions	• Time taken by the data producer entity to approve or reject all received data sharing agreements (in days)

Element Name	Element Details
	• Total number of data sharing agreements received by the producer entity
Calculation	= Time taken by the data producer entity to approve or reject all received data sharing agreements (in days) / Total number of data sharing agreements received by the producer entity
Measurement Unit	Days
Acceptable Threshold	10 Days
Scale Intervals	Unacceptable: > 10 days Low: (8 days, 10 days] Fair: (6 days, 8 days] Good: (4 days, 6 days] Excellent: (2 days, 4 days] Leader: <= 2 days
Version History	
Dependencies	

4.2.2. Open Data (OD)

Element Name	Element Details
Metric ID	OD.OE.01
Metric Name	Datasets published in ODP
Metric Description	This metric measures the percentage of datasets published by the entity in ODP against the total number of datasets required to be published in ODP by the entity. This metric aims at expediting the efforts in enriching ODP with high-impact and wide-spectrum open datasets generated by the various government entities. It also aims at increasing the contribution of the platform towards the development of innovative services, applications, and new business ideas.
Domain Name	Open Data (OD)
Data Platforms	Open Data Platform (ODP)
Definitions	• Number of datasets published in ODP by the entity. • Total number of datasets required to be published in ODP by the entity
Calculation	= Number of datasets published in ODP by the entity / Total number of datasets required to be published in ODP by the entity * 100
Measurement Unit	Percentage
Acceptable Threshold	70%
Scale Intervals	Unacceptable: <= 70% Low: (70%, 75%] Fair: (75%, 80%] Good: (80%, 85%] Excellent: (85%, 90%] Leader: > 90%

Element Name	Element Details
Version History	
Dependencies	The total number of open datasets to be published by each entity will be determined based on the size and nature of the business of that entity.

Element Name	Element Details
Metric ID	OD.OE.02
Metric Name	Delay/Lag in refreshing open datasets
Metric Description	This metric measures the delay in refreshing the published datasets in ODP by the entity. It helps entities adhere to the refresh schedule (update frequency) for the published datasets. The prime focus of this metric is to improve the data freshness by ensuring that the datasets are updated as per the expected update frequency which helps in serving the data consumers with most up-to-date data for various forms of consumptions.
Domain Name	Open Data (OD)
Data Platforms	Open Data Platform (ODP)
Definitions	• Delay in refreshing a dataset (in days) by the entity as per the update frequency of the dataset. • Update frequency of the dataset (in days) • Total number of datasets that have been published in ODP by the entity
Calculation	= Average of delay percentage for all entity's datasets, where the delay percentage of one dataset is calculated as: = Sum of the delay in refreshing the dataset (in days) by the entity as per its update frequency / (number of expected refreshes * update frequency of the dataset) * 100

Element Name	Element Details
Measurement Unit	Percentage
Acceptable Threshold	10%
Scale Intervals	Unacceptable: > 10% Low: (8%, 10%] Fair: (6%, 8%] Good: (4%, 6%] Excellent: (2%, 4%] Leader: <= 2%
Version History	
Dependencies	

Element Name	Element Details
Metric ID	OD.OE.03
Metric Name	Reported issues for the published datasets
Metric Description	This metric measures the average number of issues reported by the end users on the entity's published datasets in ODP. This metric aims at improving the overall quality of the published datasets as they have far-reaching impact on the services, research, applications, and business models developed on top of them.
Domain Name	Open Data (OD)
Data Platforms	Open Data Platform (ODP)
Definitions	• Number of issues reported on the entity's published datasets in ODP • Total number of datasets published in ODP by the entity
Calculation	= Number of issues reported on the entity's published datasets in ODP / Total number of datasets published in ODP by the entity

Element Name	Element Details
Measurement Unit	Number of issues
Acceptable Threshold	5
Scale Intervals	Unacceptable: > 5 Low: (4, 5] Fair: (3, 4] Good: (2, 3] Excellent: (1, 2] Leader: <= 1
Version History	
Dependencies	

Element Name	Element Details
Metric ID	OD.OE.04
Metric Name	Delay in resolving reported issues on published datasets
Metric Description	This metric measures the delay in resolving the issues reported on the entity's published datasets in ODP. This metric aims at optimizing the time for remediating the reported issues as the affected datasets have far-reaching impact on the services, research, applications, and business models developed on top of them.
Domain Name	Open Data (OD)
Data Platforms	Open Data Platform (ODP)
Definitions	• Time taken by the entity to resolve an issue reported on a dataset (in days) • Expected resolution time (in days)

Element Name	Element Details
Calculation	= Average of delay percentage of fixing all the issues reported on the entity's datasets, where the delay percentage associated with one issue is calculated as: = (Time taken by the entity to resolve an issue reported on a dataset – Expected resolution time of that issue) / Expected resolution time of that issue * 100
Measurement Unit	Percentage
Acceptable Threshold	10%
Scale Intervals	Unacceptable: > 10% Low: (8%, 10%] Fair: (6%, 8%] Good: (4%, 6%] Excellent: (2%, 4%] Leader: <= 2%
Version History	
Dependencies	Issues on ODP will be categorized based on their expected resolution time.

4.2.3. Data Catalog and Metadata (MCM)

Element Name	Element Details
Metric ID	MCM.OE.01
Metric Name	Systems cataloged in NDC
Metric Description	This metric measures the percentage of systems cataloged by the entity in NDC against the total number of the entity's critical systems. Note that a system is considered cataloged if its technical metadata is fully scanned and uploaded in NDC. This metric aims at accelerating the documentation of the national data assets for consumers to explore the business context of those assets with appropriate technical details, improving the data accessibility and discovery.
Domain Name	Data Catalog and Metadata (MCM)
Data Platforms	National Data Catalog (NDC)
Definitions	• Number of critical systems that have been cataloged in NDC by the entity. • Total number of the entity's critical systems
Calculation	= Number of critical systems that have been cataloged in NDC by the entity / Total number of the entity's critical systems * 100
Measurement Unit	Percentage
Acceptable Threshold	70%
Scale Intervals	Unacceptable: <= 70% Low: (70%, 75%] Fair: (75%, 80%] Good: (80%, 85%]

Element Name	Element Details
	Excellent: (85%, 90%]
	Leader: > 90%
Version History	
Dependencies	The number of critical systems that need to be cataloged by the entity will be determined based on the size and nature of the business of that entity.

Element Name	Element Details
Metric ID	MCM.OE.02
Metric Name	Business attributes defined and linked in NDC
Metric Description	This metric measures the percentage of the business attributes that are defined and linked to technical columns in NDC by the entity against the total number of business attributes that the entity is required to define in NDC. This metric aims at accelerating the documentation of the national data assets for consumers to explore the business context of those assets with appropriate technical details, improving the data accessibility and discovery.
Domain Name	Data Catalog and Metadata (MCM)
Data Platforms	National Data Catalog (NDC)
Definitions	• Number of business attributes defined and linked to technical columns in NDC by the entity • Total number of business attributes required to be defined by the entity

Element Name	Element Details
Calculation	= Number of business attributes defined and linked to technical columns in NDC by the entity / Total number of business attributes required to be defined by the entity * 100
Measurement Unit	Percentage
Acceptable Threshold	70%
Scale Intervals	Unacceptable: <= 70% Low: (70%, 75%] Fair: (75%, 80%] Good: (80%, 85%] Excellent: (85%, 90%] Leader: > 90%
Version History	
Dependencies	The total count of required business attributes for a system is determined based on the quantity and complexity of the business functions implemented in that system.

Element Name	Element Details
Metric ID	MCM.OE.03
Metric Name	Reporting assets defined in NDC
Metric Description	This metric measures the percentage of the reporting assets (KPIs/metrics) that are defined in NDC by the entity against the total number of reporting assets that the entity is required to document in NDC. This metric helps in documenting the important analytics assets which will be made discoverable for other consumers to leverage for their analysis and decision-making. This in turn promotes the

Element Name	Element Details
	reusability of those authentic and verified assets to reduce the time to market for the consumers.
Domain Name	Data Catalog and Metadata (MCM)
Data Platforms	National Data Catalog (NDC)
Definitions	• Number of reporting assets defined in NDC by entity. • Total number of required reporting assets to be defined in NDC by entity
Calculation	= Number of reporting assets defined in NDC by the entity / Total number of required reporting assets to be defined in NDC by the entity * 100
Measurement Unit	Percentage
Acceptable Threshold	70%
Scale Intervals	Unacceptable: <= 70% Low: (70%, 75%] Fair: (75%, 80%] Good: (80%, 85%] Excellent: (85%, 90%] Leader: > 90%
Version History	
Dependencies	The total number of required reporting assets to be defined is determined, upon onboarding the entity in NDC, based on the size of the entity and the number of its critical systems.

Element Name	Element Details
Metric ID	MCM.OE.04
Metric Name	Business attributes linked to attribute class standards in NDC

Element Name	Element Details
Metric Description	This metric measures the percentage of business attributes linked to attribute class standards in NDC by the entity against the total number of business attributes that are defined by that entity. This metric aims at accelerating the efforts for standardizing the national data assets to improve the overall quality of data. It will eventually help in achieving the seamless data sharing and interoperability with full trust across business processes, applications, and systems.
Domain Name	Data Catalog and Metadata (MCM)
Data Platforms	National Data Catalog (NDC)
Definitions	• Number of business attributes linked to attribute class standards in NDC by the entity • Total number of business attributes defined in NDC by the entity
Calculation	= Number of business attributes linked to attribute class standards in NDC by the entity / Total number of business attributes defined in NDC by the entity * 100
Measurement Unit	Percentage
Acceptable Threshold	70%
Scale Intervals	Unacceptable: <= 70% Low: (70%, 75%] Fair: (75%, 80%] Good: (80%, 85%] Excellent: (85%, 90%] Leader: > 90%
Version History	
Dependencies	

Element Name	Element Details
Metric ID	MCM.OE.05
Metric Name	Accuracy of business attribute relationships in NDC
Metric Description	This metric measures the percentage of issues reported on the business attributes that incorrectly linked to other metadata objects in NDC (e.g., business attributes linked to incorrect attribute class standards, incorrect technical columns or incorrect measures). This metric aims at enhancing the accuracy of the metadata which in turn increases the quality and trust of the consumers' analysis and decision-making.
Domain Name	Data Catalog and Metadata (MCM)
Data Platforms	National Data Catalog (NDC)
Definitions	• Number of incorrectly linked business attributes defined in NDC by the entity • Total number of business attributes defined in NDC by the entity
Calculation	= Number of business attributes incorrectly linked by the entity in NDC / Total number of business attributes defined in NDC by the entity * 100
Measurement Unit	Percentage
Acceptable Threshold	10%
Scale Intervals	Unacceptable: > 10% Low: (8%, 10%] Fair: (6%, 8%] Good: (4%,6%] Excellent: (2%, 4%] Leader: <= 4%
Version History	
Dependencies	

4.2.4. Reference and Master Data Management (RMD)

Element Name	Element Details
Metric ID	RMD.OE.01
Metric Name	Publishing reference entities
Metric Description	This metric measures the percentage of standardized reference entities (tables) that are published by the entity on the GSB against the number of reference entities (produced by that entity) that are required by the government entities to perform their business functions. This metric aims at accelerating the standardization of the reference values which is required by different business processes for trusted reference data interoperability.
Domain Name	Reference and Master Data Management (RMD)
Data Platforms	Government Service Bus (GSB) and Reference Data Management Platform (RDP)
Definitions	• Number of published reference entities by the entity on GSB • Total number of reference entities that are expected to be published by the entity
Calculation	= Number of published reference entities by the entity on GSB / Total number of reference entities that are expected to be published by the entity * 100
Measurement Unit	Percentage
Acceptable Threshold	90%
Scale Intervals	Unacceptable: <= 90% Low: (90%, 92%] Fair: (92%, 94%] Good: (94%, 96%]

Element Name	Element Details
	Excellent: (96%, 98%]
	Leader: > 98%
Version History	
Dependencies	

Element Name	Element Details
Metric ID	RMD.OE.02
Metric Name	Time taken to publish new reference entities
Metric Description	This metric measures the average time taken by the entity to publish new reference entities requested by other entities on GSB.

This metric aims at improving the overall availability of reference values to bring agility in building accurate business services, analysis, and reference data interoperability. |
| Domain Name | Reference and Master Data Management (RMD) |
| Data Platforms | Government Service Bus (GSB) and Reference Data Management Platform (RDP) |
| Definitions | • Total time taken by the entity to publish a reference entity on GSB (in days)
• Total number of reference entities that are published by the entity |
Calculation	= Sum of total time taken by the entity to publish its reference entities on GSB (in days) / Total number of reference entities that are published by the entity
Measurement Unit	Days
Acceptable Threshold	30 Days

Element Name	Element Details
Scale Intervals	Unacceptable: > 30 days Low: (25 days, 30 days] Fair: (20 days, 25 days] Good: (15 days, 20 days] Excellent: (10 days, 15 days] Leader: <= 10 days
Version History	
Dependencies	

Element Name	Element Details
Metric ID	RMD.OE.03
Metric Name	Time taken to fix reported issues in reference entities
Metric Description	This metric measures the average time taken by the entity to resolve an issue reported on its reference entities that are published on GSB. This metric aims at minimizing and containing the data quality issues across the government systems and applications. It also helps in maintaining the quality of the outcomes pertaining to the analysis, business processes, and data interoperability as incorrect or missing reference values will negatively impact the consumer experience and trust.
Domain Name	Reference and Master Data Management (RMD)
Data Platforms	Government Service Bus (GSB) and Reference Data Management Platform (RDP)
Definitions	• Time taken by the entity to resolve a reported issue related to its reference entities (in days)

Element Name	Element Details
	• Total number of reported issues related to the entity's reference entities published on GSB
Calculation	= Sum of time taken by the entity to resolve a reported issue related to its reference entities (in days) / Total number of reported issues related to the entity's reference entities published on GSB
Measurement Unit	Days
Acceptable Threshold	5 Days
Scale Intervals	Unacceptable: > 5 days Low: (4 days, 5 days] Fair: (3 days, 4 days] Good: (2 days, 3 days] Excellent: (1 days, 2 days] Leader: <= 1 day
Version History	
Dependencies	

4.2.5. Data Quality (DQ)

Element Name	Element Details
Metric ID	DQ.OE.01
Metric Name	Data Quality (DQ) index in GSB
Metric Description	This metric calculates the average DQ index of the entity's data that is shared on GSB by applying DQ rules that span across the various DQ dimensions on all of the entity's GSB attributes. Note that DQ rules and attributes have a many-to-many relationship. That is, a DQ rule can be applied on more than one attribute and there can be more than one rule associated with the same attribute. The notation "Rule, Attribute" will be used to refer to a pair of a DQ rule applied on an attribute. A record is considered clean with regards to a DQ rule only if the attribute on which the DQ rule is applied passes the logic implemented within that rule. This metric aims at improving the quality of the data required for business process integration and ensuring trust and seamless interoperability of data in the Kingdom.
Domain Name	Data Quality (DQ)
Data Platforms	Government Service Bus (GSB)
Definitions	• Total number of clean records for an attribute checked by a DQ rule • Total number of records checked by the DQ rule

Element Name	Element Details
Calculation	"Attribute, Rule" index = (Total number of clean records for the attribute checked by the DQ rule / Total number of records checked by the DQ rule) * 100 The "Attribute, Rule" index is subsequently rolled up using weightages applied on both attributes and DQ rules to calculate the DQ index of the entity
Measurement Unit	Percentage
Acceptable Threshold	90%
Scale Intervals	Unacceptable: <= 90% Low: (90%, 92%] Fair: (92%, 94%] Good: (94%, 96%] Excellent: (96%, 98%] Leader: > 98%
Version History	
Dependencies	

Element Name	Element Details
Metric ID	DQ.OE.02
Metric Name	Data Quality (DQ) index in NDL
Metric Description	This metric calculates the average DQ Index of the entity's data that is integrated in NDL by applying DQ rules that span across the various DQ dimensions. Since entities have a large number of attributes integrated in NDL, only critical attributes will be considered in the calculation of this index.

Element Name	Element Details
	Note that DQ rules and attributes have a many-to-many relationship. That is, a DQ rule can be applied on more than one attribute and there can be more than one rule associated with the same attribute. The notation "Rule, Attribute" will be used to refer to a pair of a DQ rule applied on an attribute. A record is considered clean with regards to a DQ rule if the attribute on which the DQ rule is applied passes the logic implemented within that rule. This metric aims at improving the quality of the data within NDL to serve the consumers with high fidelity data for their decision-making needs. It also supports in identifying and prioritizing the key areas of focus for improving the overall quality of the national data assets.
Domain Name	Data Quality (DQ)
Data Platforms	National Data Lake (NDL)
Definitions	• Total number of clean records for an attribute checked by a DQ rule • Total number of records checked by the DQ rule
Calculation	"Attribute, Rule" index = (Total number of clean records for the attribute checked by the DQ rule / Total number of records checked by the DQ rule) * 100 The "Attribute, Rule" index is subsequently rolled up using weightages applied on both attributes and DQ rules to calculate the DQ index of the entity
Measurement Unit	Percentage
Acceptable Threshold	70%

Element Name	Element Details
Scale Intervals	Unacceptable: <= 70%
	Low: (70%, 75%]
	Fair: (75%, 80%]
	Good: (80%, 85%]
	Excellent: (85%, 90%]
	Leader: > 90%
Version History	
Dependencies	

4.2.6. Data Operations (DO)

Element Name	Element Details
Metric ID	DO.OE.01
Metric Name	Delay in response time of GSB APIs
Metric Description	This metric measures the delay in the response time of the entity's GSB APIs after subtracting the time taken by the GSB platform. This metric only considers the actual processing time incurred by the entity after receiving a request from the consumer. This metric aims at improving the business process integration and interoperability, and hence enhancing the consumer experience and application resilience.
Domain Name	Data Operations (DO)
Data Platforms	Government Service Bus (GSB)
Definitions	• Response time of an API which is the amount of time (in milliseconds) taken by the API to process a request and generate a response • Expected response time of the API (in milliseconds) • Number of calls for an API
Calculation	= Average of delay percentage for all entity's APIs, where the delay percentage of one API is calculated as: = Sum of delay in response time of all API's calls / (Number of calls for an API * expected response time of the API) * 100 where the delay in response time of an API call is calculated as follows:

Element Name	Element Details
	= Response time of the API – GSB latency – expected response time of the API
Measurement Unit	Percentage
Acceptable Threshold	10%
Scale Intervals	Unacceptable: > 10% Low: (8%, 10%] Fair: (6%, 8%] Good: (4%, 6%] Excellent: (2%, 4%] Leader: <= 2%
Version History	
Dependencies	The expected response time of an API will be determined based on its processing complexity and payload size. Prior to measuring this metric, all APIs of each entity will need to be categorized into a defined set of categories in order to determine their expected response time.

Element Name	Element Details
Metric ID	DO.OE.02
Metric Name	Failed API calls on GSB
Metric Description	This metric measures the rate of failed API calls invoked by the consumer applications on the entity's GSB APIs where the application receives an error of any type from the called API.

Element Name	Element Details
	This metric aims at improving the resilience of the GSB APIs and hence improving the business processes efficiency and the consumer experience.
Domain Name	Data Operations (DO)
Data Platforms	Government Service Bus (GSB)
Definitions	• Number of failed API calls • Number of API calls
Calculation	= Number of failed API calls / Number of API calls * 100
Measurement Unit	Percentage
Acceptable Threshold	1%
Scale Intervals	Unacceptable: > 1% Low: (0.8%, 1%] Fair: (0.6%, 0.8%] Good: (0.4%, 0.6%] Excellent: (0.2%, 0.4%] Leader: <= 0.2%
Version History	
Dependencies	

Element Name	Element Details
Metric ID	DO.OE.03
Metric Name	Operational issues from entities encountered by NDL
Metric Description	This metric measures the percentage of operational issues encountered by NDL from each source entity against the number of data pipeline executions with the respective entity. This metric will be measured by NDL for each source entity with which NDL has data pipelines operated.

Element Name	Element Details
	The prime focus of this metric is to improve the data freshness by reducing the recurring errors and issues which will help in serving the consumers with most up-to-date data for various forms of consumptions.
Domain Name	Data Operations (DO)
Data Platforms	National Data Lake (NDL)
Definitions	• Number of failed pipeline executions caused by operational issues from the entity. • Total number of data pipeline executions with the entity
Calculation	= Number of failed pipeline executions caused by operational issues from the entity / Total number of data pipeline executions with the entity * 100
Measurement Unit	Percentage
Acceptable Threshold	6%
Scale Intervals	Unacceptable: > 6% Low: (5%, 6%] Fair: (4%, 5%] Good: (3%, 4%] Excellent: (2%, 3%] Leader: <= 2%
Version History	
Dependencies	

5 Measurement Approach

Each government entity will be measured against these metrics on an annual basis. Each year/round will have a different prioritization based on various factors such as the importance, the availability of required metadata, and the criticality of the metrics, and these factors will be translated into the weights of the metrics.

6 References

[1]. National Data Management Office – Data Management and Personal Data

Protection Standards (https://sdaia.gov.sa/ndmo/Files/PoliciesEn001.pdf)