Republic of Zambia
Office of the President
Electronic Government Division – SMART Zambia Institute

PUBLIC SERVICE INFORMATION COMMUNICATION TECHNOLOGY STANDARDS

DATA INTEROPERABILITY STANDARDS

Version: 1.0
Author: Electronic Government Division - Standards Department
Document Classification: Public

Published Date: January 2024

Document History:

Validation and Distribution

	Name	Issue date
Issued by	The Electronic Government Division – SMART Zambia Institute	2024
Verified by	Standards Task Team	December 2023
Approved by	National Coordinator	December 2023

Distribution List	
1	Cabinet Office
2	All Ministries and Spending Agencies
3	Online publishing (SZI Website)

Document Revision History:

Version	Date	Author	Remarks
1.0	2024	Standards and Compliance Department	Creation of Documents

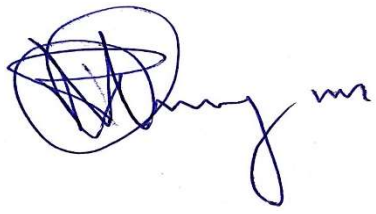
FOREWORD

The Electronic Government Division, SMART Zambia Institute is mandated with formulating and enforcing standards in Information and Communication Technologies (ICT) across all Ministries, Provinces and Spending Agencies (MPSAs) to facilitate the nation's transition into a Digital Society. In view of its mandate, the e-Government Division has developed the Data Interoperability Standard that is meant to facilitate the seamless exchange, integration, and use of data across different Government departments, agencies, and systems.

The Data Interoperability Standard has been issued to MPSAs to promote interoperability ensuring that Government Systems and databases work together, even if they were developed independently. This will enhance collaboration among Government institutions by standardizing data formats, definitions and exchange protocols providing a holistic appraisal to addressing complex siloed data systems.

The implementation of this standard will be monitored by the National Electronic Government Council while the e-Government Division will undertake the enforcement of this document. Annual audits shall be carried out in all MPSAs to determine their levels of compliance to this standard. The e-Government Division shall issue a certificate of compliance to an MPSA upon completion of a successful audit assessment. For non-compliant MPSAs, a report detailing the extent of the deviation and the prevailing circumstances shall be tabled before the National Electronic Government Council who will advise on the appropriate action to be taken.

All MPSAs are required to ensure full compliance to the Data Interoperability Standard for effective and efficient public service delivery.

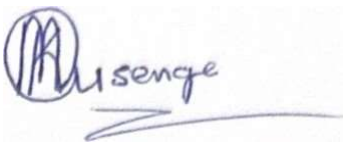


Percy Chinyama (Mr.)
National Coordinator
Electronic Government Division

ACKNOWLEDGEMENT

The development of the Data Interoperability Standard addresses a multitude of critical elements, including data quality, seamless data sharing, regulatory compliance, data-driven decision-making, and the harmonious integration of data across the spectrum of Government entities. It is poised to elevate Governments capacity for service delivery, enhance the accuracy and consistency of our data, and foster innovative solutions to the complex challenges we face.

It is for this reason that I wish to express my sincere gratitude to the e-Government Standards Task Team for their unwavering dedication to the development of this standard. Their expertise, tireless efforts, and visionary leadership have been instrumental in the development of this standard. The division is also indebted to the Heads of ICT in Ministries, Provinces, and Spending Agencies (MPSA) and stakeholders, whose guidance and commitment have propelled this initiative forward. This document will ensure standardisation of critical data elements providing seamless exchange of information in Government.



Kasali Musenge
Director Standards and Compliance
Electronic Government Division

Contents

Working Definitions

Data Element (1)	The smallest unit of a data structure, e.g., a column of a table; a dataset separated by XML tags. Also called "data unit."
Data Element (2)	A unit of data for which the definition, identification, representation, and permissible values are specified by means of a set of attributes.
Data Element Name	A single or multi-word designation used as the primary means of identification of data elements for humans.
Data Structure	A physical or logical relationship among units of data and the data themselves.
Database:	A structured collection of data organized according to a conceptual structure describing the characteristics of these data and the relationships among their corresponding entities, supporting one or more application areas. Also, a data structure for accepting, storing, and providing on-demand data for multiple independent users.
Definition:	A representation of a concept by a descriptive statement that serves to differentiate it from related concepts.
Domain	A field of special knowledge.
Domain Expert:	A specialist in a field of special knowledge.
Domain Glossary:	An alphabetical list of terms with definitions, commonly containing explanations of words, concepts, or terms in alphabetical order within a particular field or subject matter.
End User:	A person/role using an operation and usually also the semantic description of this operation.
ERD (Entity Relationship Diagram)	A model of the relationship between a database and tables.

Glossary	An alphabetical or systematic list of all words in a language or words of a certain category, including explanations in that language or translations into one or more languages.
GML (Geospatial Markup Language):	A language for marking up geospatial information.
Human-Readable Description	A description in a form readable to both developers and domain experts, disclosing the meaning of data (semantics).
IDABC (Interoperable Delivery of European eGovernment Services to public Administrations, Businesses, and Citizens)	An initiative aimed at delivering interoperable eGovernment services.
IDEF (Integration Definition)	A family of modelling languages in the field of systems and software engineering.
IETF (Internet Engineering Task Force):	A community that develops and promotes Internet standards.
ITU-T (International Telecommunication Union for Telecommunication Standards):	A specialized agency of the United Nations that develops and publishes telecommunications standards.
Machine-Readable Description:	A description in a form readable to software systems that conveys the data without changing their meaning (semantics).
Name:	A term used for referring to a domain concept (human-readable).
OASIS XCBF (OASIS XML Common Biometric Format):	A standard format for representing biometric data in XML.
Object Class	A set of objects classified or grouped on the basis of a common property.
Object Status	A status experienced during the object's life cycle.

OMG (Object Management Group):	A consortium that develops various standards, including the UML (Unified Modelling Language).
Ontology	A way of describing a domain in both machine- and human-readable forms. Includes a glossary of concepts and relations used in the particular domain to create an agreed-upon vocabulary for exchanging information.
Ontology Description Language	A way of describing a domain in both machine- and human-readable forms, including a glossary of concepts and relations used in the particular domain.
OWL (Ontology Web Language)	A markup language for presenting ontologies on the World Wide Web.
Process:	A certain operation or sequence of operations of an organization necessary for fulfilling a task or providing a service. Also called "operational process" or "business process."
RDF (Resource Description Framework):	A framework for describing resources on the web.
RDFS (Resource Description Framework Schema):	A markup language for presenting ontologies on the World Wide Web, similar to OWL.
Recommendation	A requirement that is not mandatory.
RSS (Really Simple Syndication):	A format for delivering regularly changing web content.
Rule:	A mandatory requirement to be followed.
SAML (Security Assertion Markup Language):	An XML-based standard for exchanging authentication and authorization data.
SA-WSDL (Semantic Annotations for WSDL and XML Schema)	A W3C recommendation for the semantic description of web services and data structures.

Semantic Web Standards	A common framework that allows data to be shared and reused across application, enterprise, and community boundaries.
UML (Unified Modelling Language)	A language for modelling and visually presenting software systems and business domains.
Value Domain	A set of permissible values, often used in the context of ISO/IEC 11179
W3C (World Wide Web Consortium)	The main organization developing web standards.
WSDL (Web Services Description Language)	An XML-based language for describing web services.
XML, or eXtensible Markup Language	Extensible Markup Language is a markup language and file format for storing, transmitting, and reconstructing arbitrary data. It defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.
XMI (XML Metadata Interchange)	A standard allowing the exchange of UML models between parties in a standard way.
XSL (Extensible Stylesheet Language):	A language for transforming XML documents into other formats.

DATA INTEROPERABILITY STANDARDS

1. INTRODUCTION

As digitalisation of public services takes root in the public sector and the call for interoperability becomes paramount, standardization of data in information systems is essential to modern governance and economic development. The ability to access, share, and utilise data efficiently and harmoniously across diverse Information Communication Technology (ICT) systems is fundamental to achieving the full potential of digital transformation. As our nation embraces the digital age, the need for a comprehensive ICT Data Interoperability Standard has never been more critical.

This Standard outlines the framework, guidelines, and best practices for achieving data interoperability within the ICT ecosystem of our country. By setting clear standards and expectations for data exchange, integration, and compatibility, this standard aims to break down silos, facilitate collaboration, and unlock the full value of our nation's data resources. It is a testament to our commitment to harnessing the power of data for the betterment of our citizens, businesses, and the overall socio-economic landscape.

This document further presents the principles, components and implementation strategies that underpin the Public Sector ICT Data Interoperability Standard. This standard addresses the complexities of modern data systems, promotes transparency and safeguards data security. This standard will serve as a guide for all stakeholders in our ICT ecosystem, ensuring that data interoperability becomes a reality, leading to a more connected, efficient, and innovative digital future for Zambia.

1.1. Background

Prior to the establishment of the Electronic Government Division, each Ministry, Province and Spending Agency (MPSA) independently embarked on the acquisition and development of information systems, a practice that inadvertently led to the proliferation of siloed systems. Within this framework, each entity opted for its own set of data standards. Over the years, this approach has resulted in the creation of numerous isolated systems within and across public institutions, each designed to collect, process, and disseminate data for the purposes of service delivery and national development. Concurrent with this expansion, the emergence of diverse technology platforms, data definitions, and institutional arrangements has underscored the growing need to allocate resources for the integration of data that underpins effective policy formulation and decision-making.

At its core, interoperability signifies the ability to harmoniously connect and amalgamate data without compromising its intrinsic meaning. In practical terms, data is considered interoperable when it can be effortlessly reused and processed across different applications, allowing disparate information systems to seamlessly collaborate. Achieving this demands the adherence to universally identifiable data standards. Notably, interoperability stands as a pivotal driver in the Government's transition towards a more data-driven approach.

In today's digital landscape, the populace anticipates heightened interconnectivity and seamless interoperability, where diverse systems can efficiently deliver data to those who require it, tailored to their specific needs and in a consistent and predictable manner. Consequently, data interoperability and integration have emerged as pivotal components of data management strategies within the Government. Nonetheless, public service institutions often find themselves immersed in the demands of day-to-day operations, leaving limited time and resources for the introduction and adoption of standards, technologies, tools, and best practices necessary for achieving enhanced data interoperability.

The absence of a Data Interoperability standard has, in various instances, compelled the Government to allocate substantial resources to the development of complex application programming interfaces (API's) for the integration of siloed systems. Furthermore, manual interventions have been necessitated to transform data into formats suitable for consumption by other Government systems.

In light of these challenges, the quest for a comprehensive Data Interoperability Standard has become an imperative, representing a fundamental shift toward a more efficient, connected, and data-responsive Government. As such, thus standard is envisioned to streamline data management, foster innovation, and ultimately enhance the quality of public services to the citizenry.

1.2. Purpose

The purpose of the Data Interoperability Standards are;

1. To provide a framework for improving data sharing by defining universally understandable concepts and standards for data-sharing communities, reducing the need for complex mediations.
2. To ensure compliance with the minimum level of interoperability within public service institutions and
3. To serve as guideline for Data formatting and structures in information systems.

1.3. Scope

This Data Interoperability standard shall apply to:

1. Information system developers, including architects, analysts, designers, programmers, record-keepers, project managers, and other relevant roles.
2. Database owners, users, integrators of operations, and contracting entities.
3. Information system owners, system auditors, and relevant stakeholders in both the public and private sectors.
4. Government agencies directly involved in data acquisition, solution architecture, data design, and implementation.

2. SEMANTIC INTEROPERABILITY

Semantic interoperability encompasses:

(a) The ability of organizations to understand exchanged data in a consistent manner.
(b) The capacity of software systems to effectively utilise data received from other systems.
(c) The establishment of relationships between data structures and real-world objects, relations, and events.
(d) The exchange of contextual information about data, including relations, operations, and general functioning.
(e) The exchange of metadata between organizations/agencies.

2.1. Achieving Semantic Interoperability

In order for Semantic interoperability to be achieve, the following are key:

(a) Data exchange partners share a mutual understanding of the meaning of the shared data.
(b) Data exchanges align with the shared understanding.
(c) Data is exchanged without ambiguity or errors.

2.2. Types of Incompatibilities

Semantic interoperability shall address three primary categories of incompatibilities:

(a) Domain Level Incompatibilities: These arise from conflicts in naming, data representation, and data scaling.
    i. Naming Conflicts: Conflicts that arise when different concepts are described using the same word, when the same attribute name holds different meanings (homonyms), when multiple alternative words describe the same concepts, or when different attribute names refer to the same thing (synonyms).
    ii. Representation Conflicts: Arising from representing values in different ways, such as using different representations for the same value type (e.g., representing sex as "M" and "F" in one system and "male" and "female" in another).
    iii. Spatial Domain Conflict: Arising when there are different legal implications for the same piece of data, for example, "blood group" information being allowed in one context but violating privacy rights in another.
(b) Entity Level Incompatibilities: These result from conflicts in measurement, confounding, schema isomorphism, and structural conflicts.

i. Measurement Conflicts: Arising from data being represented in different units or scales, such as differences in length units like "meters" and "feet."
ii. Confounding Conflicts: Resulting from assigning different meanings to a single concept, for example, using different interpretations of stock prices.
iii. Schema Isomorphism Conflicts: Arising when the same concept is described with a different set of attributes.
iv. Structural Conflicts: Arising when the same piece of information is modelled as a relation name, an attribute name, or a value in a table.

(c) Abstraction Level Incompatibilities: These occur when semantically similar entities or attributes are represented at different levels of abstraction, including integrity, generalisation, computational, aggregation, and granularity conflicts.
i. Integrity Conflicts: Resulting from data considered correct in one context violating integrity constraints in another context.
ii. Generalization Conflicts: Arising when data in one context may be a subset or superset of another, leading to mandatory attributes in one context being invalid in another.
iii. Computational Conflicts: Resulting from alternative methods of computing data.
iv. Aggregation Conflicts: Arising when data stored in one context is defined collectively in another (or vice versa).
v. Granularity Conflicts: Occurring when data is reported at different levels of abstraction or granularity.

2.3. Benefits of Semantic Interoperability

The advantages of semantic interoperability include:

(a) Enhanced data quality and accuracy, reducing errors and discrepancies resulting from misinterpretation or duplication of data.
(b) Improved data integration, enabling more effective decision-making through a shared understanding of data.

3. DATA INTEROPERABILITY

3.1. Data Interoperability

Data interoperability is essential for the seamless exchange of data across different systems and organizational boundaries. It includes:

(a) The ability to correctly interpret data across different systems or organizational boundaries.
(b) Facilitating a common understanding of data meaning and usage between systems and across agencies – providing clarity in plain English or familiar business language.

3.2. Implementing Data Interoperability

Implementing data interoperability requires achieving both data integration and data exchange as well as enabling effective use of the data that becomes available. Each of these tasks involves some type of standards and guidelines in the way data is captured and consumed between disparate systems. Data semantics irregularities are most commonly evidenced through differences in:

(a) Data names;
(b) Data types;
(c) Data lengths; and
(d) Data structures

3.3. Standards & Principal Statement Relating to Data Interoperability

(a) Use Extensible Markup Language (XML) and XML Schema for Data Interoperability;
(b) Use XML Metadata Interchange (XMI) for the exchange of all business information and information system modelling;
(c) Use Extensible Stylesheet Language (XSL) for data transformation;
(d) Ensure XML products are written to comply with the recommendations of the World Wide Web Consortium (W3C);
(e) Where necessary, base the work on the draft W3C standards but avoid the use of any product-specific XML extensions that are not being considered for open standardization with the W3C.

3.4. Recommended Standards and Specifications

Recommended standards and specifications for Data Interoperability and Transformation are as stipulated in Appendix 1; while Standards for biometric interchange are as stipulated in Appendix 2.

3.5. Data Element Definitions

Data element definitions shall be the primary means for conveying data meaning, ensuring seamless data exchange, and defining agreements between parties. Guidelines for standardizing data definitions are in Appendix B.

4. STANDARD FOR DATA ELEMENT DEFINITIONS

4.1. Data Element Definitions

i. Data element definitions shall be the primary vehicle for conveying the meaning of data.
ii. Data element definitions shall, in addition to being precise and unambiguous, ensure a seamless exchange of data.
iii. When two or more parties exchange data, the meaning of that data shall be explicit in an agreement.
iv. Definitions shall be written to facilitate understanding by any user and by recipients of shared data.

4.2. Benefits of Standardizing Data Definition

Some of the benefits of standardizing data definition include:

i. Standardize structure and contents of metadata registries;
ii. Make metadata collections accessible, searchable by semantic content;
iii. Support understanding and reuse of data standards; and
iv. Promote use of standards for greater interoperability.

4.3. Guidelines for Data Definition

i. A data element definition must be standardized using a common structure, including metadata.
ii. Metadata shall include a data element identifier, name, definition, synonyms, usage notes, data type, format, and other relevant attributes.
iii. Definitions shall be presented in a clear, concise, and human-readable manner.
iv. Definitions shall be easily accessible and searchable.\\
v. Definitions shall include appropriate context to aid understanding.

4.4. Standardized Data Element Definitions

i. Data element definitions shall follow the guidelines in Appendix 3.
ii. Data element definitions shall be published in a central repository for access and use by relevant stakeholders.

5. DATA EXCHANGE STANDARDS

5.1. Data Exchange Standards

i. Data exchange standards shall be used to facilitate the exchange of data between different systems, organizations, and agencies.

ii. Data exchange standards shall define the format, structure, and rules governing the transfer of data.

5.2. Implementing Data Exchange Standards

i. Implementing data exchange standards requires adherence to the defined standards and guidelines.
ii. Data exchange standards shall be developed, maintained, and published for access by relevant stakeholders.
iii. Data exchange standards shall ensure the accurate and secure transmission of data.

5.3. Benefits of Data Exchange Standards

The benefits of data exchange standards include:

i. Improved data consistency and accuracy;
ii. Reduced data transfer errors;
iii. Enhanced security and privacy;
iv. Streamlined data integration and sharing.

5.4. Guidelines for Data Exchange Standards

i. Data exchange standards shall define the format and structure of data messages.
ii. Data exchange standards shall include data validation rules and error handling procedures.
iii. Data exchange standards shall support data encryption and security measures.
iv. Data exchange standards shall be well-documented and readily available to stakeholders.

5.5. Recommended Data Exchange Standards

Recommended data exchange standards are as stipulated in Appendix 4.

6. METADATA STANDARDS

6.1. Metadata Standards

i. Metadata standards shall be used to describe and document data element definitions, data exchange standards, and other relevant information.
ii. Metadata standards shall provide a common framework for organizing and managing metadata.

6.2. Implementing Metadata Standards

i. Implementing metadata standards requires adherence to the defined standards and guidelines.
ii. Metadata standards shall define the structure and elements of metadata records.
iii. Metadata standards shall support the discovery, understanding, and use of data and information resources.

6.3. Benefits of Metadata Standards

The benefits of metadata standards include:

i. Improved data discovery and access;
ii. Enhanced data documentation and understanding;
iii. Facilitated data management and governance;
iv. Increased data interoperability and usability.

6.4. Guidelines for Metadata Standards

i. Metadata standards shall define metadata elements and their attributes.
ii. Metadata standards shall include guidelines for metadata creation, maintenance, and dissemination.
iii. Metadata standards shall support machine-readable formats and interoperability.

6.5. Recommended Metadata Standards

Recommended metadata standards are as stipulated in Appendix 6.

7. DATA SECURITY AND PRIVACY

Standards for Data Security and Privacy shall be as provided for in the Public Service information Security Standard.

7.1. Data Security and Privacy

i. Data security and privacy standards shall be implemented to protect sensitive and confidential information.
ii. Data security and privacy standards shall define safeguards and controls to prevent unauthorized access, disclosure, or misuse of data.

7.2. Implementing Data Security and Privacy Standards

i. Implementing data security and privacy standards requires adherence to the defined standards and guidelines.

ii. Data security and privacy standards shall be developed, maintained, and enforced to ensure the protection of data.

7.3. Benefits of Data Security and Privacy Standards

The benefits of data security and privacy standards include:

i. Protection of sensitive information from unauthorized access or breaches;
ii. Compliance with legal and regulatory requirements;
iii. Enhanced trust and confidence in data sharing and exchange;
iv. Mitigation of data-related risks and liabilities.

7.4. Guidelines for Data Security and Privacy Standards

i. Data security and privacy standards shall define access controls, authentication, and authorization mechanisms.
ii. Data security and privacy standards shall specify encryption and data protection measures.
iii. Data security and privacy standards shall address data retention and disposal requirements.
iv. Data security and privacy standards shall comply with applicable laws and regulations.

7.5. Recommended Data Security and Privacy Standards

Recommended data security and privacy standards are as stipulated in Appendix 7.

8. CONCLUSION

These Public Service ICT Data Interoperability Standards provide a framework for achieving semantic interoperability, data interoperability, data element definitions, data exchange standards, metadata standards, and data security and privacy. Adherence to these standards in alignment to the existing legal framework for data exchange will facilitate data sharing, integration, and exchange among different systems, organizations, and agencies in the public service, leading to improved data quality, accuracy, and usability.

9. APPENDICES

Appendix 1: Recommended Standards and Specifications for Data Interoperability and Transformation

Table 1 - Recommended Standards & Specifications for Data Interoperability and Transformation

COMPONENT	STANDARD	STANDARD BODY
Metadata/MetaLanguage	XML (Extensible Markup Language)	W3C
XML Metadata Definition	XML-Schema RelaxNG	W3C OASIS/ISO
XML Data Transformation	XSL (Extensible Stylesheet Language)	W3C
XML Data Transformation	Xpath	W3C
XML Signature	XML DSIGW3C	
XML Security Mark-up	SAML v2.0 (Security Assertion Markup Language)	OASIS
Public Key Infrastructure	X509v3 (SSL and TSL)	ITU-T
Model exchange	XMI (XML Metadata Interchange)	OMG

Appendix 2: Standards for Biometric Interchange

Table 2 - Standards for Biometric Interchange

COMPONENT	STANDARD	STANDARD BODY
Secure XML Encoding for exchanging Biometric data	OASIS XCBF 1.1 Specification	OASIS
Secure XML encodings for the patron formats specified in CBEFF (Common Biometric Exchange File Format (NISTRI 6529)	OASIS	
Data Element Specification	ISO/IEC 19785-1:2006	ISO/IEC
Information Technology – Common Biometric Exchange Formats Framework – Part 1: Data Element Specification	ISO/IEC	

Interchange Format Framework	ISO/IEC 19794: Information Technology Biometric data interchange formats – Part 1: Framework	ISO/IEC
Interchange Format Framework for finger minutiae data	ISO/IEC 19794: Information Technology Biometric data interchange formats – Part 2: Finger minutiae data	ISO/IEC
Interchange Format Framework for finger pattern spectral	ISO/IEC 19794: Information Technology Biometric data interchange formats – Part 3: Finger pattern spectral	ISO/IEC
Interchange Format Framework for finger image data	ISO/IEC 19794: Information Technology Biometric data interchange formats – Part 4: Finger image data	ISO/IEC
Interchange Format Framework for Face image data	ISO/IEC 19794: Information Technology Biometric data interchange formats – Part 5: Face image data	ISO/IEC
Interchange Format Framework for Signature/sign behaviour data	ISO/IEC 19794: Information Technology Biometric data interchange formats – Part 7: Signature/sign behaviour data	ISO/IEC
Graphical/still image information exchange specification	ISO/IEC 10918-1:1994 Information Technology – Digital compression and coding of continuous-tone still images: Requirements and Guidelines	ISO/IEC
ISO/IEC 10918-1: 1994/CD Cor 1		ISO/IEC
ISO/IEC 10918-2: 1995 Information Technology - Digital compression and coding of continuous-tone still images: Compliance testing	ISO/IEC	
ISO/IEC 10918-3: 1997 Information Technology - Digital compression and coding of continuous-tone still images: Extensions	ISO/IEC	
ISO/IEC 10918-3: 1997/Amd 1: 1999		ISO/IEC

ISO/IEC 10918-4: 1999 Information Technology - Digital compression and coding of continuous-tone still images: Registration of JPEG profiles, SPIFF profiles, SPIFF tags, SPIFF compression types and Registration Authorities (REGAUT)	ISO/IEC	

NOTE: The standards provided in Section 4.2 above apply to exchange of biometric data only and not how the data is captured. However, data sharing, and exchange will be governed by the provisions of Data Protection Act or any relevant legislation.

Appendix 3: Guidelines for Data Element Definitions

Table 3 - Guidelines for Data Element Definitions - Requirements

Requirement	Data Definition	Explanation	Example	Good Definition	Poor Definition	Reason
1	A data definition shall: be stated in the singular	The concept expressed by the data definition shall be expressed in the singular. (An exception is made if the concept itself is plural.)	“Article Number”	A reference number that identifies an article.	Reference number identifying articles.	The poor definition uses the plural word “articles,” which is ambiguous, since it could imply that an “article number” refers to more than one article.
2	A data definition shall: state what the concept is, not only what it is not	When constructing definitions, the concept cannot be defined exclusively by stating what the concept is not.	“Freight Cost Amount”	Cost amount incurred by a shipper in moving goods from one place to another.	Costs which are not related to packing, documentation, loading, unloading, and insurance.	The poor definition does not specify what is included in the meaning of the data.

3	A data definition shall: be stated as a descriptive phrase or sentence(s) (in most languages)	A phrase is necessary (in most languages) to form a precise definition that includes the essential characteristics of the concept. Simply stating one or more synonym(s) is insufficient. Simply	“Agent Name”	Cost amoName of party authorized to act on behalf of another party.unt incurred by a shipper in moving goods from one place to another.	Representative.	“Representative” is a near-synonym of the data element name, which is not adequate for a definition.

4	A data definition shall: contain only commonly understood abbreviations	Understanding the meaning of an abbreviation, including acronyms and initialisms, is usually confined to a certain environment. In other environments the same abbreviation can cause misinterpretation or confusion. Therefore, to avoid ambiguity, full words, not abbreviations, shall be used in the definition. Exceptions to this requirement may be made if an abbreviation is commonly understood such as “i.e.” and “e.g.” or if an abbreviation is more readily understood than the full form of a complex term and has	Example 1: “Tide Height”	The vertical distance from mean sea level (MSL) to a specific tide level.	The vertical distance from MSL to a specific tide level.	The poor definition is unclear because the acronym, MSL, is not commonly understood and some users may need to refer to other sources to determine what it represents. Without the full word, finding the term in a glossary may be difficult or impossible.

		been adopted as a term in its own right such as “radar” standing for “radio detecting and ranging.” All acronyms must be expanded on the first occurrence.	Example 2: “Unit of Density Measurement”	The unit employed in measuring the concentration of matter in terms of mass per unit (m.p.u.) volume (e.g., pound per cubic foot; kilogram per cubic meter).	The unit employed in measuring the concentration of matter in terms of m.p.u. volume (e.g., pound per cubic foot; kilogram per cubic meter).	m.p.u. is not a common abbreviation, and its meaning may not be understood by some users. The abbreviation should be expanded to full words.

5	A data definition shall: be expressed without embedding definitions of other data or underlying concepts	As shown in the following example, the definition of a second data element or related concept should not appear in the definition proper of the primary data element. Definitions of terms should be provided in an associated glossary. If the second definition is necessary, it may be attached by a note at the end of the primary definition's main text or as a separate entry in the dictionary. Related definitions can be accessed through relational attributes (e.g., cross-reference).	Example 1: “Sample Type Code”	A code identifying the kind of sample.	A code identifying the kind of sample collected. A sample is a small specimen taken for testing. It can be either an actual sample for testing, or a quality control surrogate sample. A quality control sample is a surrogate sample taken to verify results of actual samples.	The poor definition contains two extraneous definitions embedded in it. They are definitions of “sample” and of “quality control sample.”

			Example 2: "Issuing Bank Documentary Credit Number"	Reference number assigned by issuing bank to a documentary credit.	Reference number assigned by issuing bank to a documentary credit. A documentary credit is a document in which a bank states that it has issued a documentary credit under which the beneficiary is to obtain payment, acceptance, or negotiation on compliance with certain terms and conditions and against presentation of stipulated documents and such drafts as may be specified.	The poor definition contains a concept definition, which should be included in a glossary.

Table 4 - Guidelines for Data Element Definitions- Recommendations

Recommendations	Data Definition	Explanation	Example	Good Definition	Poor Definition	Reason
1	A data definition should: state the essential meaning of the concept	All primary characteristics of the concept represented should appear in the definition at the relevant level of specificity for the context. The inclusion of non-essential characteristics should be avoided. The level of detail necessary is dependent upon the needs of the system user and environment.	Example 1: "Consignment Loading Sequence Number" (Intended context: any form of transportation)	A number indicating the sequence in which consignments are loaded in a means of transport or piece of transport equipment.	A number indicating the sequence in which consignments are loaded in a truck.	In the intended context, consignments can be transported by various transportation modes, e.g., trucks, vessels or freight trains. Consignments are not limited to trucks for transport.

			Example 2: “Invoice Amount”	Total sum charged on an invoice.	The total sum of all chargeable items mentioned on an invoice, taking into account deductions on the one hand, such as allowances and discounts, and additions on the other hand, such as charges for insurance, transport, handling, etc.	The poor definition includes extraneous material.

2	A data definition should: be precise and unambiguous	The exact meaning and interpretation of the defined concept should be apparent from the definition. A definition should be clear enough to allow only one possible interpretation.	“Shipment Receipt Date”	Date on which a shipment is received by the receiving party.	Date on which a specific shipment is delivered.	The poor definition does not specify what determines a “delivery.” “Delivery” could be understood as either the act of unloading a product at the intended destination or the point at which the intended customer actually obtains the product. It is possible that the intended customer never receives the product that has been unloaded at his site or the customer may receive the product days after it was unloaded at the site.

3	A data definition should: be concise	The definition should be brief and comprehensive. Extraneous qualifying phrases such as “for the purpose of this metadata registry,” “terms to be described,” shall be avoided.	“Character Set Name”	The name given to the set of phonetic or ideographic symbols in which data is encoded.	The name given to the set of phonetic or ideographic symbols in which data is encoded, for the purpose of this metadata registry, or, as used elsewhere, the capability of systems hardware and software to process data encoded in one or more scripts.	In the poor definition, all the phrases after “...data is encoded” are extraneous qualifying phrases.
4	A data definition should: be able to stand alone	The meaning of the concept should be apparent from the definition. Additional explanations or references should not be necessary for understanding the meaning of the definition.	“School Location City Name”	Name of the city where a school is situated.	See “school site”.	The poor definition does not stand alone, it requires the aid of a second definition (school site) to understand the meaning of the first.

5	A data definition should: be expressed without embedding rationale, functional usage, domain information, or procedural information	Although they are often necessary, such statements do not belong in the definition proper because they contain information extraneous to the definition. If deemed useful, such expressions may be placed in other metadata attributes (see ISO/IEC 11179-3). It is, however, permissible to add examples after the definition.	“Data Field Label”	Identification of a field in an index, thesaurus, query, database, etc.	Identification of a field in an index, thesaurus, query, database, etc., which is provided for units of information such as abstracts, columns within tables.	The poor definition contains remarks about functional usage. This information starting with “which is provided for...” must be excluded from the definition and placed in another attribute, if it is necessary information.
		1. The rationale for a given definition should not be included as part of the definition (e.g. if a data element uses miles instead of kilometres, the reason should not be indicated in the definition).				
		2. Functional usage such as: “this data element should not be used for ...” should not be included in the definition proper.				
		3. Remarks about procedural aspects. For example, “This data element is used in conjunction with data element 'xxx'”, should not appear in the definition; instead use “Related data reference” and “Type of relationship” as specified in ISO/IEC 11179-3.				

6	A data definition should: avoid circular reasoning	Two definitions shall not be defined in terms of each other. A definition should not use another concept's definition as its definition. This results in a situation where a concept is defined with the aid of another concept that is, in turn, defined with the aid of the given concept.	two data elements with poor definitions: 1. 1) Employee ID Number - Number assigned to an employee. 2. 2) Employee - Person corresponding to the employee ID number.			Each definition refers to the other for its meaning. The meaning is not given in either definition.
7	A data definition should: use the same terminology and consistent logical structure for related definitions	A common terminology and syntax should be used for similar or associated definitions.	The following example illustrates this idea. Both definitions pertain to related concepts and therefore have the same logical structure and similar terminology.			Using the same terminology and syntax facilitates understanding. Otherwise, users wonder whether some difference is implied by use of synonymous terms and variable syntax.

			1. “Goods Dispatch Date” - Date on which goods were dispatched by a given party.			
			2. “Goods Receipt Date” - Date on which goods were received by a given party.			

Appendix 4: Recommended Data Exchange Standards

Table 5 – Logical Data Element Names

Prime Word	Qualifier (Secondary Prime Word)	Class Word	Logical Name
Account		Balance	Account Balance
Employee	Salary	Amount	Employee Salary Amount
Student	Last	Name	Student Last Name

Table 6 - Physical Data Element

Logical Name	DB 1	DB 2
Account Balance	ACCT_BAL	ACCT-BAL
Employee Salary Amount	EMP_SAL_AMT	EMP-SAL-AMT
Student Last Name	STU_LST_NAME	STU-LST-NAME

Appendix 6: Recommended Metadata Standards

Note that the Domain column fields identified as “Text” can be any domain of text data, char, varchar, nvarchar or other. Also numeric is any number domain as int, double and other.

Table 7 - Recommended Metadata Standards

Suffix Classwords	Description	Domain	Total Length	Decimal Places
Address	Descriptive text used to denote a place where a person or organization may be communicated with (i.e., PO Box), or a physical location (i.e., Street Address).	Text, nvarchar	60	
Amount	Most values of the number data type expressed as two decimal places which are meant to define currency. For special accounting purposes four digits to the right of the decimal point are acceptable.	Number	15	4
BLOB	A subset of Blobs that are non-imaged data. This includes data that is compressed, zipped, or otherwise encrypted	Binary	Variable	
Code	A numeric or character value that identifies classifications or categories of a member of a set of like values. A code does not include the description of the code value rather a simple abbreviation that stands for that description.	Number or Text as is required	Integer or variable not exceeding 9 bytes	
Constant	Data which does not change value over time or in different circumstances or uses.	Text or Number	Numeric or Variable not exceeding 8 bytes	
Date	A unit of time expressed in months, days, and years, used to designate a specific 24-hour period	Date	8	
Datetime	A specific instance of time that includes date and time components	Dependent on the DBMS		

		used		
Decimal	A numeric representation of data that is not normally considered a quantity and represented in float decimal or numeric notation with or without significant digits to the right of the decimal point	Number preferred Or Float if no other options is available	Variable	Variable
Description	Data having undefined, freeform, unstructured, or unformatted content and is not an Address or Name	Text, nvarchar	Variable	
Flag	A bit or series of bits with two stable states. A binary condition permitting only two values (i.e., True/False, Yes/No, Pass/Fail)	Char(1) or bit	1	
Hash	A resulting hash from Secure Hash Algorithm-256 (SHA256) or higher when available	Text, nvarchar	Variable	
ID	Either a numeric value that implies sequence; or a computer-generated serial identifier used to generate primary keys in a database to maintain referential integrity	Number	9	
Image	A subset of Binary Large Objects, Blobs, that represents a digitized or scanned image or document. (This includes PDF's, bitmaps, jpegs and other image forms and document types	Binary	Variable	
Name	A word or phrase that constitutes the distinctive designation of a person, place, thing, or concept	Text, nvarchar	Variable	
ShortName	A word or phrase that constitutes the distinctive designation of a person, place, thing, or concept in a shortened string. Acronyms, codes, etc. will be valid uses of this field	Text, nvarchar	Variable	

Number	A combination of letters and/or numbers used to uniquely identify an occurrence of something. (i.e., Social Security Number, Vehicle Identification Number). Special characters used as separators would be excluded from occurrences of attributes or fields in this class. Rather, display formats would achieve this effect. For example, Social Security Number would be 9 digits without the 2 „-"separators	Number or Text as is required	Variable	
Percent	A number that represents the ratio between two values that have the same unit of measure multiplied by 100 (A rate times 100)	Number	5	2
Quantity	A number of non-monetary units expressed in conjunction with a unit of measure	Number	Dependent on associated Measurement and Unit	
Rate	A quantity or amount measured with respect to another measured quantity or amount (i.e., naira/hour, miles/gallon, etc.)	Number	9	4
XML	A valid XML Document, which could be XML, XSLT, Schema, or other well-formed XML document type	XML	Variable	