



Intelligence Community Technical Specification

XML Data Encoding Specification for DigitalHazMat Assertion

Version 2021-JAN

January 15, 2021

Distribution Notice:

This document has been approved for Public Release and is available for use without restriction.

Table of Contents

Chapter 1 - Introduction	1
1.1 - Purpose	1
1.2 - Scope	1
1.3 - Enterprise Need	1
1.4 - Conventions	1
1.4.1 - XML Namespaces	2
1.5 - Dependencies	2
1.5.1 - Specification Dependencies	2
1.5.2 - Inverse Dependencies	4
Chapter 2 - Development Guidance	6
2.1 - Understanding DigitalHazMat Assertion	6
2.2 - DHZM Assertion Usage	6
Chapter 3 - Constraints	7
3.1 - Data Validation Constraint Rules	7
3.1.1 - Purpose	7
3.1.2 - Value Enumeration Constraints	7
3.1.3 - Additional Constraints	7
3.1.3.1 - DES Constraints	7
3.1.4 - Constraint Rules	7
3.2 - Data Rendering Constraint Rules	7
3.2.1 - Purpose	7
3.2.2 - Rendering Constraint Rules	8
Appendix A - Feature Summary	9
A.1 - DHZM Feature Summary	9
Appendix B - Change History	10
B.1 - V2021-JAN Initial Release Summary	10
Appendix C - Glossary	11
Appendix D - List of Abbreviations	12
Appendix E - Bibliography	14
Appendix F - Points of Contact	16
Appendix G - IC CIO Approval Memo	17

List of Figures

Figure 1 - Related Specifications 4

Figure 2 - Inverse Dependency Specifications 5

List of Tables

Table 1 - XML Namepaces	2
Table 2 - Direct Dependencies	3
Table 3 - Constraint Rules	8
Table 4 - Feature Summary Legend	9
Table 5 - DHZM Feature comparison	9
Table 6 - DES Version Identifier History	10
Table 7 - Data Encoding Specification V2021-JAN Initial Release Summary	10

Chapter 1 - Introduction

1.1 - Purpose

This *XML Data Encoding Specification for DigitalHazMat Assertion* (DHZM.XML) defines detailed implementation guidance for using Extensible Markup Language (XML) to encode DHZM data. This Data Encoding Specification (DES) defines the XML elements and attributes, associated structures and relationships, mandatory and cardinality requirements, and permissible values for representing DHZM data assertion concepts using XML within the use of a Trusted Data Format (TDF) Object.

1.2 - Scope

The *Intelligence Community Technical Specification Framework* (IC-SF.XML^[2]) defines the basic conceptual structure and outlines the core philosophy of Intelligence Community (IC) technical specifications. For convenience, a copy of this framework is included in every package.

This specification is applicable to the IC and information produced by, stored, or shared within the IC. This DES may have relevance outside the scope of intelligence; however, prior to applying outside of this defined scope, the DES should be closely scrutinized and differences separately documented and assessed for applicability.

1.3 - Enterprise Need

This DES is designed to fulfill a number of requirements in support of the transformational efforts of the IC. These requirements include:

- the need for a shared common technique to securely encapsulate, obfuscate and transport potentially malicious artifacts from one network to another.
- the need to normalize metadata to facilitate communication between groups (e.g., government, commercial industry) with regards to digital hazmat provenance and analysis (e.g., isolated environments for triage, classified environments for deep analysis, multiple assertions for group analysis).

Both enterprise needs and requirements for this specification can be found in the following policies and implementation guidance:

- 500 Series:
 - Intelligence Community Directive (ICD) 500, *Director Of National Intelligence Chief Information Officer*^[3]
 - ICD 501, *Discovery and Dissemination or Retrieval of Information within the IC*^[4]
 - Intelligence Community Standard (ICS) 500-20, *IC Enterprise Standards Compliance*^[5]

1.4 - Conventions

Certain technical and presentation conventions are used in the creation of the IC technical specifications to ensure readability and understanding. For details, please see the "Specifications Conventions" chapter in the IC-SF.XML^[2].

1.4.1 - XML Namespaces

Namespaces referenced in this document and the prefixes used to represent them are listed in the following table. The namespace prefix of any XML Qualified Name used in any example in this document should be interpreted using the information below.

Table 1 - XML Namespaces

Prefix	URI
dhzm	urn:us:gov:ic:digitalhazmat
tdf	urn:us:gov:ic:tdf

1.5 - Dependencies

Specifications often rely on other specifications, components or artifacts, either directly or indirectly. For specific definitions of dependency terminology used throughout this section, please see the “Dependency Definitions” chapter in the IC-SF.XML^[2].

1.5.1 - Specification Dependencies

This technical specification directly depends on the technical specifications, documentation, and implementations listed in [Table 2](#). The dependencies listed below are directly referenced in this specification (e.g., Schema, Schematron), and are normative or informative as indicated.

The subsequent figure, [Figure 1](#), is an informative graphical representation of all of the Intelligence Community Chief Information Officer (IC CIO) specifications related to this specification. The graphic depicts dependencies. However, the representations may not match an exact schema import tree or dependency diagram that an analysis of the Schema, Schematron or other documents would yield. For example, the graphic only shows a given specification once even though it may actually be imported by many specifications or be a direct dependency. All specifications listed in [Table 2](#) will be shown in [Figure 1](#); however not all specifications listed in [Figure 1](#) may appear in [Table 2](#). [Figure 1](#) is to aid users in gaining a general understanding of all dependencies.

Table 2 - Direct Dependencies

Name	Dependency Description
<i>XML Data Encoding Specification for DigitalHazMat Commercial TDF</i> (DHZMC-TDF.XML.V2021-JAN+ ^[6])	DHZM.XML elements as well as its dependent specifications are used in conjunction with DHZMC-TDF.XML ^[6] objects as structured assertions or content that compose the necessary material represented by DHZM.XML. The dependence of DHZM.XML on DHZMC-TDF is normative. This specification does not depend on a specific version of DHZMC-TDF.XML; DHZMC-TDF.XML versions later than version 2021-JAN MAY be used. The minimum version was based on a technical dependency.
<i>Intelligence Community Specification Framework</i> (IC-SF.XML.V2021-JAN+ ^[2])	This specification does not depend on a specific version of IC-SF.XML ^[2] ; versions later than version 2021-JAN MAY be used, however, the newest version of IC-SF.XML SHOULD be used as IC-SF.XML is expected to always replace its preceding version. The minimum version was based on technical dependencies on IC-SF.XML; IC-SF.XML is the basic structure of and philosophy behind intelligence community technical specifications.
Schematron ^[8]	<p>Schematron — International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) 19757-3:2006 — is a rule-based document schema definition language. In this specification Schematron is a formal language used to express normative business rules, so this reference is normative.</p> <p>The Schematron rules are normative in the sense that they convey criteria that a document MUST adhere to, exactly as English may be used to convey normative criteria. It is not necessary for implementers to use the specific Schematron encoding in this specification. Implementers MAY use any encodings, tools, or languages desired to implement validation schemes for conformance to this specification.</p> <p>Note: The Schematron rules in this specification use Transformations (XSLT) 2.0^[9] query binding.</p>

Name	Dependency Description
<p>XSLT 2.0^[9] implementation of Schematron^[8] by Rick Jelliffe (2010-04-14)</p> <p>Note: The only available identifying descriptors for this implementation are the implementer's name and date of release. This implementation may be found at the following Uniform Resource Locator (URL): http://code.google.com/p/schematron/.</p>	<p>The International Organization for Standardization does not create nor endorse reference implementations of its standards. For the purposes of this specification the <i>behavior</i> of the implementation created by Mr. Jelliffe is normative.</p> <p>Implementers MAY use any encodings, tools, or languages desired to implement validation schemes for conformance to this specification. To conform to this specification, a validator MUST find a document valid <i>if and only if</i> the Schematron implementation by Mr. Jelliffe would find the document valid according to the Schematron rules in this specification.</p>

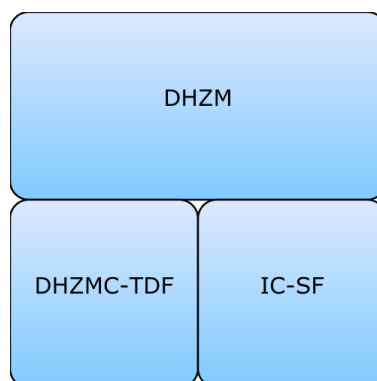


Figure 1 : Related Specifications

1.5.2 - Inverse Dependencies

Generally, it is only necessary to think of the *direct dependencies* in the dependency tree. However, with the specification versions being decoupled, it is also important to consider the *inverse dependencies*, for compatibility with newer versions of a given specification. The changes introduced to a given specification can sometimes make it incompatible with current versions of its inverse dependencies.

Since this specification is one such specification that is used by other specifications released by the IC CIO, the [Figure 2](#) has been included to assist readers in understanding all of the dependency relationships and how changes in a specification may impact others. This diagram is representative of dependencies at the time of the release of this specification, but are subject to change over time.

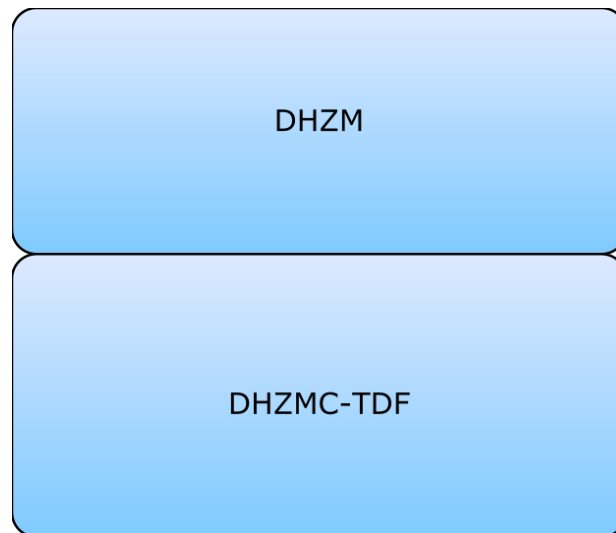


Figure 2 : Inverse Dependency Specifications

Chapter 2 - Development Guidance

For information on the structure and content of the specifications, please see the “Specification Overview” chapter in the IC-SF.XML^[2] framework document. This chapter is intended to expand upon the common information that the framework specifies providing specific development guidance that is specific to the implementation of this specification.

2.1 - Understanding DigitalHazMat Assertion

The encoding of a digital hazmat assertion is made up of at least one of the following components:

- A **@dhzm:ProvenanceAssertion** contains details about the origin and history of the digital hazmat.
- A **@dhzm:AnalysisAssertion** contains information that is used in the analysis of the digital hazmat and also results of the analysis fo the digital hazmat.

2.2 - DHZM Assertion Usage

DHZM.XML is used in conjunction with DHZMC-TDF.XML^[6] and exists as structured assertions. A Trusted Data Object (TDO) conforms to DHZM.XML when it contains at least one of the following.

- A structured assertion of **@tdf:scope="PAYL"** and a **dhzm:ProvenanceAssertion** element
- A structured assertion of **@tdf:scope="PAYL"** and a **dhzm:AnalysisAssertion** element

Chapter 3 - Constraints

3.1 - Data Validation Constraint Rules

3.1.1 - Purpose

The DHZM.XML schema defines the data elements, attributes, cardinalities and parent-child relationships for which XML instances must comply. Validation of these syntax aspects is an important first step in the validation process. An additional level of validation is needed to ensure that the content complies with the constraints as specified in applicable IC policy guidance and codified in these constraint rules. Traditional schema languages are generally unable to effectively represent these additional constraints. For more information, please see the “Data Validation Constraint Rules” chapter in the IC-SF.XML^[2] framework document.

3.1.2 - Value Enumeration Constraints

DHZM.XML currently does not contain any Controlled Vocabulary Enumeration (CVE)s.

3.1.3 - Additional Constraints

3.1.3.1 - DES Constraints

The DES version is specified through attributes on the root element. The schema constrains the values of these attributes. The `@DESVersion` attribute enables systems processing an instance document to be certain which set of constraint rules, schema, CVEs and business rules are intended by the author to be used.

3.1.4 - Constraint Rules

The detailed constraint rules for the DHZM.XML schema can be found in a separate document inside the Documents/DHZM directory, in the “DHZM_Rules.pdf” file. This document is generated from the individual Schematron files to provide a single searchable document for all of the constraint rules encoded in Schematron. Obsolete rule numbers are listed in the “DHZM_Rules.pdf” file.

3.2 - Data Rendering Constraint Rules

3.2.1 - Purpose

Rendering rules define constraints on the rendering and display of DHZM.XML documents. The intent is to inform the development of systems capable of rendering or displaying DHZM.XML data for use by individuals not familiar with the details of the DHZM.XML markup. While expressed in a similar manner to the data validation constraint rules above, there is no expectation that evaluation of these rules can be automated; rather these rules should inform the evaluation of a system’s capabilities and functionality.

3.2.2 - Rendering Constraint Rules

The following table contains the information for the DHZM.XML data rendering constraint rules.

Table 3 - Constraint Rules

Rule Number	Severity	Description	Human Readable Description
There are no Data Rendering Constraint rules at this time.			

Appendix A Feature Summary

The following tables summarize major features by version for DHZM.XML. The “Required date” is the date when systems SHOULD support a feature based on the specified driver. Executive Orders, Information Security Oversight Office (ISOO) notices, ICDs and other policy documents have a variety of effective dates. The “Required date” may be later than the date of applicable policy based on the effective date defined in the policy (e.g., The IC Marking System Register and Manual^[1] has an implementation date of one year after issuance).

Table 4 - Feature Summary Legend

Key	Description
F	Full (able to comply and verified by spec to some degree)
P	Partial (Able to comply but not verifiable)
N	Non-compliance (Can’t comply)
N/A	Not Applicable. Feature is no longer required.
Cell Colors represent the same information as the Key value	

A.1. DHZM Feature Summary

Table 5 - DHZM Feature comparison

Required date	Feature	V2021-JAN
	Defines the allowable values for DigitalHazMat Assertion	F

Appendix B Change History

The following table summarizes the version identifier history for this DES.

Table 6 - DES Version Identifier History

Version	Date	Purpose
2021-JAN	January 15, 2021	Initial Release. For details, see Section B.1 - V2021-JAN Initial Release Summary

B.1 - V2021-JAN Initial Release Summary

Significant drivers for Version V2021-JAN include:

- Creation of DHZM.XML specification.

The following table summarizes the initial release in V2021-JAN.

Table 7 - Data Encoding Specification V2021-JAN Initial Release Summary

#	Change	Artifacts changed	Compatibility Notes
1	Creation of DHZM.XML specification. (CR-2019-060, CR-2020-020, CR-2020-021, CR-2020-022, CR-2020-023, CR-2020-024, CR-2020-025, CR-2020-026, CR-2020-027, CR-2020-028, CR-2020-029, CR-2020-030, CR-2020-035, CR-2020-038, CR-2020-039, CR-2020-041)	Documentation Schema Schematron	Initial Release.

Appendix C Glossary

This appendix lists terms, definitions and sources of the definitions for terms used in this document.

Entity	<p>An individual (person), organization, device, or process.</p> <p>Source: NIST 800-56Br1, <i>Recommendation for Pair-Wise Key-Establishment Schemes Using Integer Factorization Cryptography, Revision 1</i> [7].</p>
NMTOKEN	<p>A built-in XML datatype based on the token datatype. The datatype NMTOKEN represents a single string token. NMTOKEN values may consist of letters, digits, periods (.), hyphens (-), underscores (_) and colons (:). No whitespace may appear within an NMTOKEN.</p> <p>Source: https://www.w3.org/TR/2004/REC-xmlschema-2-20041028/</p>
token	<p>A token datatype is an XML schema language built-in datatype. A token datatype is a string datatype that contains one or more strings separated by a single space, e.g., <code>ism:releasableTo='USA AFG FVEY'</code> is an example of an ISM attribute that has token datatype. A token datatype contains no leading or trailing spaces, no carriage returns, no line feeds and no tab characters. The individual strings in an element or attribute that is a token datatype are referred to as <u>tokens</u>. In the <code>ism:releasableTo='USA AFG FVEY'</code> example, the tokens are 'USA', 'AFG' and 'FVEY'. In contrast, the <u>value</u> of <code>ism:releasableTo</code> is the entire string 'USA AFG FVEY'.</p> <p>Source: https://www.w3.org/TR/2004/REC-xmlschema-2-20041028/#token</p>

Appendix D List of Abbreviations

This appendix lists all the acronyms and abbreviations referenced in this encoding specification.

AAS	Authoritative Attribute Source
AS	Attribute Service
CVE	Controlled Vocabulary Enumeration
DES	Data Encoding Specification
DNI	Director of National Intelligence
IC	Intelligence Community
ICAM	Identity, Credential, and Access Management
IC ICAM SC	IC Identity, Credential, and Access Management Subcommittee
IC CIO	Intelligence Community Chief Information Officer
IC IE	IC Information Environment
ICD	Intelligence Community Directive
ICS	Intelligence Community Standard
IEC	International Electrotechnical Commission
ISM	Information Security Markings
ISO	International Organization for Standardization
ISOO	Information Security Oversight Office
NPE	Non-Person Entity
OCIO	Office of the Intelligence Community Chief Information Officer
PE	Person Entity
TDC	Trusted Data Collection
TDF	Trusted Data Format
TDO	Trusted Data Object
URL	Uniform Resource Locator
XML	Extensible Markup Language
XSL	Extensible Stylesheet Language

XSLT

XSL Transformations

Appendix E Bibliography

Bibliography

[1] IC Markings

Director of National Intelligence (DNI), Special Security Directorate (SSD), Security Markings Program (SMP). *Intelligence Community Markings System Register and Manual*. Available online Intelink-TS at: <https://go.ic.gov/tGXkwGO>
Available online Intelink-U at: <https://w3id.org/ic/standards/policy/icmarkings>

[2] IC-SF.XML

Office of the Director of National Intelligence. *Intelligence Community Specification Framework (IC-SF.XML)*. Available online Intelink-TS at: <https://go.ic.gov/pNFyuVg>
Available online Intelink-U at: <https://w3id.org/ic/standards/IC-SF>
Available online at: <https://w3id.org/ic/standards/public>

[3] ICD 500

Office of the Director of National Intelligence. *Director of National Intelligence Chief Information Officer*. Intelligence Community Directive 500. 7 August 2008. Available online Intelink-TS at: <https://go.ic.gov/U7v6ZRL>
Available online at: http://www.dni.gov/files/documents/ICD/ICD_500.pdf

[4] ICD 501

Office of the Director of National Intelligence. *Discovery and Dissemination or Retrieval of Information within the Intelligence Community*. Intelligence Community Directive 501. 21 January 2009. Available online Intelink-TS at: <https://go.ic.gov/fTBM8OS>
Available online at: http://www.dni.gov/files/documents/ICD/ICD_501.pdf

[5] ICS 500-20

Director of National Intelligence Chief Information Officer. *Intelligence Community Enterprise Standards Compliance*. Intelligence Community Standard 500-20. 16 December 2010. Available online Intelink-TS at: <https://go.ic.gov/kh8NMVJ>
Available online Intelink-U at: <https://w3id.org/ic/standards/policy/ICS500-20>

[6] DHZMC-TDF.XML

Office of the Director of National Intelligence. *XML Data Encoding Specification for DigitalHazMat Commercial TDF (DHZMC-TDF.XML)*. Available online Intelink-TS at: <https://go.ic.gov/UPDATEME>
Available online Intelink-U at: <https://w3id.org/ic/standards/DigitalHazMatCommercialTDF>
Available online at: <https://w3id.org/ic/standards/public>

[7] NIST 800-56Br1

National Institute of Standards and Technology. *Recommendation for Pair-Wise Key-Establishment Schemes Using Integer Factorization Cryptography*. Revision 1. September 2014.

Available online at: <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-56Br1.pdf>

[8] Schematron

International Organization for Standardization (ISO). *Information technology -- Document Schema Definition Language (DSDL) -- Part 3: Rule-based validation -- Schematron*. ISO/IEC 19757-3:2006.

ISO Spec Available online at: <http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html>

StyleSheets for compiling Available online at: <http://code.google.com/p/schematron/>

[9] XSLT2

World Wide Web Consortium (W3C). *XSL Transformations (XSLT) Version 2.0*. W3C Recommendation 23 January 2007.

Available online at: <http://www.w3.org/TR/xslt20/>

Bibliography

Appendix F Points of Contact

The Intelligence Community Chief Information Officer (IC CIO) facilitates one or more collaboration and coordination forums charged with the adoption, modification, development, and governance of IC technical specifications of common concern. This technical specification was produced by the IC CIO and coordinated with these forums, approved by the IC CIO or a designated representative, and made available at the following Director of National Intelligence (DNI)-sponsored web sites.

Public Website: <https://w3id.org/ic/standards/public>

Intelshare: <https://w3id.org/ic/standards/data-specs>

Direct all inquiries about this IC technical specification, IC technical specification collaboration and coordination forums, or IC element representatives involved in those forums, to the IC CIO.

E-mail: ic-standards-support@odni.gov.

Appendix G IC CIO Approval Memo

An Office of the Intelligence Community Chief Information Officer (OCIO) Approval Memo should accompany this enterprise technical data specification bearing the signature of the IC CIO or an IC CIO-designated official(s). If an OCIO Approval Memo is not accompanying this specification's version release package, then refer back to the authoritative web location(s) for this specification to see if a more complete package or a specification update is available.

Specification artifacts display a date representing the last time a version's artifacts as a whole were modified. This date most often represents the conclusion of the IC Element collaboration and coordination process. Once the IC Element coordination process is complete, the specification goes through an internal OCIO staffing and coordination process leading to signature of the OCIO Approval Memo. The signature date of the OCIO Approval Memo will be later than the last modified date shown on the specification artifacts by an indeterminable time period.

Upon signature of the OCIO Approval Memo, IC Elements may begin to use this specification version in order to address mission and business objectives. However, it is critical for IC Elements, prior to disseminating information encoded with this new specification version, to ensure that key enterprise services and consumers are prepared to accept this information. IC Elements should work with enterprise service providers and consumers to orchestrate an orderly implementation transition to this specification version in concert with mandatory and retirement usage decisions captured in the IC Enterprise Standards Baseline as defined in ICS 500-20^[5].