

XML DATA ENCODING SPECIFICATION FOR INFORMATION RESOURCE METADATA VERSION 3

ICTechSpec 500.D.5-V3

An Intelligence Community Technical Specification
Prepared by the
Intelligence Community Chief Information Officer

6 December 2010



Table of Contents

| | |
|---|-----------|
| Table of Contents | i |
| List of Tables | ii |
| Chapter 1 – Introduction | 2 |
| 1.1 Purpose | 2 |
| 1.2 Needs and Requirements | 2 |
| 1.3 Audience and Applicability..... | 2 |
| 1.4 Utility..... | 3 |
| 1.5 Version Information | 3 |
| 1.6 Components of this DES | 3 |
| 1.7 Normative and Informative Components | 4 |
| 1.8 Technical Encoding Dependencies | 4 |
| 1.9 Typographic conventions | 4 |
| Chapter 2 – Development Guidance | 5 |
| 2.1 Mapping of Abstract Data Elements to Physical XML Elements..... | 5 |
| 2.2 Additional guidance | 10 |
| 2.2.1 <i>DescribedItemIRM & MetaCardIRM</i> | 10 |
| 2.2.2 <i>DocumentID</i> | 11 |
| 2.2.3 <i>ISM Attribute usage</i> | 11 |
| 2.2.4 <i>Specification of publishing organization</i> | 11 |
| 2.2.5 <i>MIME type</i> | 13 |
| Chapter 3 – XML Schema Guide | 15 |
| Chapter 4 – Constraint Rules | 16 |
| 4.1 Basics | 16 |
| 4.1.1 <i>Classified or Controlled Constraint Rules</i> | 16 |
| 4.1.2 <i>Terminology</i> | 16 |
| 4.1.3 <i>Rule Identifiers</i> | 17 |
| 4.1.4 <i>Errors and Warnings</i> | 17 |
| 4.2 Global Constraints | 17 |
| 4.3 DES Constraints | 17 |
| 4.4 Value Enumeration Constraints..... | 18 |
| 4.5 General Constraints | 19 |
| 4.6 Information Security Markings (ISM.XML) | 20 |
| Appendix A IC CIO Approval Memo..... | 22 |
| Appendix B Acronyms..... | 23 |
| Appendix C Glossary..... | 25 |
| Appendix D Bibliography | 26 |
| Appendix E Points of Contact | 28 |
| Appendix F Change History | 29 |
| F.1 V3 Change Summary | 30 |
| F.2 V2 Change Summary | 31 |

| | | |
|-------------------|---|-----------|
| Appendix G | Configuration Management | 32 |
| Appendix H | Controlled Vocabulary Enumerations | 33 |

List of Tables

| | |
|---|----|
| Table 1. Mapping of Abstract Data Elements to Physical XML Elements | 6 |
| Table 2. DES Version Identifier History | 29 |
| Table 3. Data Encoding Specification V3 Change Summary | 30 |
| Table 4. Data Encoding Specification V2 Change Summary | 31 |

Chapter 1 – Introduction

1.1 Purpose

This *XML Data Encoding Specification for Information Resource Metadata* (IRM.XML) defines detailed specifications for using Extensible Markup Language (XML) to encode information resource metadata in compliance with the *Intelligence Community (IC) Abstract Data Definition*. This Data Encoding Specification (DES) defines the XML elements and attributes, associated structures and relationships, mandatory and cardinality requirements, and permissible values for representing security marking concepts using XML.

This DES uses the Department of Defense Discovery Metadata Specification (DDMS) as a base and builds on that base by specifying additional metadata needed to describe information resources in the IC. In some cases, this DES specifies additional constraints on the data.

1.2 Needs and Requirements

Information sharing within the national intelligence enterprise will increasingly rely on information resource metadata to allow users and systems to find and access a wide-range of information resources throughout the enterprise. Information resource visibility, accessibility, and understandability are all critical to providing these capabilities. A successful information sharing enterprise depends on the ability of users and systems to locate and access information resources through a consistent and flexible search, or discovery capability. An enterprise-wide discovery capability will be greatly enhanced by the consistent “digital” description of all information resources. A common specification for the description of information resources allows for a comprehensive capability that can locate all resources across the enterprise regardless of format, type, location, or classification.

1.3 Audience and Applicability

DESS are intended primarily to be used by those developing tools and services to create, modify, store, exchange, search, display, or further process the type of data being described. The applicability and conditions for when the DES should be used will be found in the Intelligence Community Enterprise Standards Baseline and referenced in IC policy guidance.

1.4 Utility

A DES specifies how to implement the abstract concepts defined in the IC Abstract Data Definition Set (ADD) in a particular physical form (e.g., data or file format). For example:

- DESs for textual markup formats, such as XML and HyperText Markup Language (HTML), define markup elements and attributes, their relationships, cardinalities, processing requirements, and use.
- DESs for display formats, such as text and Adobe Portable Document Format (PDF), define text and typographic conventions, cardinalities, processing requirements, and use.
- DESs for application-specific formats, such as Microsoft Word, define document properties, styles, fields, cardinalities, processing requirements, and use.

1.5 Version Information

This is **Version 3** of this DES. This version number must be specified in any instance document intended to be valid against this version using the attribute **DESVersion**. A separate **DESVersion** attribute should be specified for each DES an instance document is claiming compliance with; each of these attributes should be in the namespace specified by each DES.

For descriptions of the changes made in this and prior versions see Appendix F.

1.6 Components of this DES

This document is the primary documentary component of the DES. This document contains:

- **Chapter 1 – Introduction.** The introduction describes high-level background information for this document. It defines the purpose and scope of this document.
- **Chapter 2 – Development Guidance.** This chapter covers two primary topics:
 - 1) Mappings of the XML element and attributes defined within this DES to appropriate IC.ADD data elements
 - 2) Descriptions of how particular encoding situations should be handled using the features provided by this DES.

- **Chapter 3 – XML Schema Guide.** Highlights the availability of an interactive presentation of the IRM.XML schema as well as an implementation-specific data element dictionary.
- **Chapter 4 – Data Validation Constraint Rules.** The constraint rules in this chapter define data validation constraints for IRM.XML beyond those in the XML Schema.

This DES consists of a number of additional technical components to include: the interactive *XML Schema Guide* referenced in Chapter 3, XML schema files, and Controlled Vocabulary Enumerations (CVE) files.

1.7 Normative and Informative Components

The XML schemas, CVE values from the XML CVE files, and the Chapter 4 constraint rules are normative for this DES. The rest of this document, the descriptive content referenced within the XML Schema Guide, and HTML CVE value files are informative.

1.8 Technical Encoding Dependencies

This DES relies on:

- Information Security Marking (ISM.XML.V5)
- Need To Know (NTK.XML.V3)
- Value enumerations used for several XML structures are defined in the various CVEs included in this DES.

1.9 Typographic conventions

Certain typography is used throughout the body of this document to convey certain meanings, in particular:

- *Italics* – A title of a referenced work or a specialized or emphasized term.
- Underscore – An abstract data element.
- **Bold** – An XML element or attribute.

Chapter 2 – Development Guidance

This chapter covers two primary topics:

- 1) Mappings of the XML element and attributes defined within this DES to appropriate IC.ADD data elements
- 2) Descriptions of how particular encoding situations should be handled using the features provided by this DES

2.1 Mapping of Abstract Data Elements to Physical XML Elements

The mapping of abstract data elements from the IC.ADD to the corresponding physical XML structures defined by this DES is shown in the following tables, which reflect the groupings in the IC.ADD. These mappings are provided for reference only. The complete set of DES artifacts, both normative and informative, should be consulted.

This mapping and additional mappings in other DESs provide a starting point for the development of automated transformations between formats defined by the DESs. However, it should be noted that when these transformations are used between formats with different levels of detail, there might be some data loss.

Note: DescribedItemIRM nodes pertain to the described resource (e.g. `irm:ICResourceMetadataPackage/irm:DescribedItemIRM/ddms:Resource/ddms:creator` is the creator of the document). The XPaths listed under **DescribedItemIRM nodes** are relative to `irm:ICResourceMetadataPackage/irm:DescribedItemIRM`.

MetaCardIRM nodes pertain to the metacard, rather than the described resource (e.g. `irm:ICResourceMetadataPackage/irm:MetaCardIRM/irm:CreatorPublishList/ddms:creator` refers to the creator of the metacard, not necessarily the creator of the document). The XPaths listed under **MetaCardIRM nodes** are relative to `irm:ICResourceMetadataPackage/irm:MetaCardIRM`. See section 2.2.1 for a further discussion of these concepts.

Table 1. Mapping of Abstract Data Elements to Physical XML Elements

| Abstract Data Element | Definition | XPath and XML implementation notes |
|-----------------------|---|---|
| Contributor | An entity responsible for making contributions to the resource. Examples of Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity. | DescribedItemIRM nodes ./ddms:Resource/ddms:contributor MetaCardIRM nodes ./irm:CreatorPublishList/ddms:contributor |
| Coverage | The spatial, temporal [or virtual] topic of the resource, the spatial [or virtual] applicability of the resource, or the jurisdiction under which the resource is relevant. Spatial topic may be a named place or a location specified by its geographic coordinates. Temporal period may be a named period, date, or date range. Virtual topic may be a named place or a location specified using a network or email address. A jurisdiction may be a named administrative entity or a geographic place to which the resource applies. Recommended best practice is to use a controlled vocabulary such as the Thesaurus of Geographic Names (TGN) or the NGA Geographic Names Server (GNS) as sanctioned by the United States Board on Geographic Names. Where appropriate, named places or time periods can be used in preference to numeric identifiers such as sets of coordinates or date ranges. | DescribedItemIRM nodes ./ddms:Resource/ddms:geospatialCoverage ./ddms:Resource/ddms:temporalCoverage ./ddms:Resource/ddms:virtualCoverage MetaCardIRM nodes Not applicable |
| Creator | An entity primarily responsible for making the resource. Examples of Creator include a person, an organization, or a service. Typically, the name of a creator should be used to indicate the entity. | DescribedItemIRM nodes ./ddms:Resource/ddms:creator MetaCardIRM nodes ./irm:CreatorPublishList/ddms:creator |

| Abstract Data Element | Definition | XPath and XML implementation notes |
|-----------------------|--|--|
| Date | A point or period of time associated with an event in the lifecycle of the resource. Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3CDTF profile of ISO 8601. Typically, date will be associated with the creation or availability of the resource. | <p>DescribedItemIRM nodes</p> <p>./ddms:Resource/ddms:dates/@ddms:created ./ddms:Resource/ddms:dates/@ddms:infoCutOff ./ddms:Resource/ddms:dates/@ddms:posted ./ddms:Resource/ddms:dates/@ddms:validTil ./irm:Dates/@irm:dateApproved</p> <p>MetaCardIRM nodes</p> <p>./irm:Dates/@irm:dateApproved ./ddms:dates/@ddms:created ./ddms:dates/@ddms:infoCutOff ./ddms:dates/@ddms:posted ./ddms:dates/@ddms:validTil</p> |
| Description | An account of the resource. Description may include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource. | <p>DescribedItemIRM nodes</p> <p>./ddms:Resource/ddms:description ./TaskingInfo ./Activity ./NoticeList ./ProductionMetricsList</p> <p>MetaCardIRM nodes</p> <p>./ddms:description ./irm:NoticeList</p> |
| Format | The file format, physical medium, or dimensions of the resource. Examples of dimensions include size and duration. Recommended best practice is to use a controlled vocabulary such as the list of Internet Media Types (MIME). Format may be used to identify the software, hardware, or other equipment needed to display or operate the resource. | <p>DescribedItemIRM nodes</p> <p>./ddms:Resource/ddms:format ./irm:ApplicationSoftware</p> <p>MetaCardIRM nodes</p> <p>./irm:ApplicationSoftware</p> |
| Identifier | An unambiguous reference to the resource within a given context. Recommended best practice is to identify the resource by means of a string conforming to a formal identification system. Formal identification systems include but are not limited to the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI), and the International Standard Book Number (ISBN). | <p>DescribedItemIRM nodes</p> <p>./ddms:Resource/ddms:identifier</p> <p>MetaCardIRM nodes</p> <p>./irm:IdentifierList/ddms:identifier</p> |

| Abstract Data Element | Definition | XPath and XML implementation notes |
|-----------------------|--|--|
| Language | A language of the resource. Recommended best practice is to use a controlled vocabulary such as RFC 3066, <i>Tags for the Identification of Languages</i> , which specifies use of ISO 639-2, <i>Codes for the Representation of Names of Languages</i> , three character language code, with an optional appended ISO 3166-1, <i>Codes for the representation of names of countries and their subdivisions</i> , two character country code. For example: "eng-US" or "eng-UK." | DescribedItemIRM nodes ./ddms:Resource/ddms:language MetaCardIRM nodes Not applicable |
| Publisher | An entity responsible for making the resource available. Examples of a Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity. | DescribedItemIRM nodes ./ddms:Resource/ddms:publisher ./irm:PublishingAgency MetaCardIRM nodes ./irm:CreatorPublishList/ddms:publisher ./irm:PublishingAgency |
| Relation | A related resource. Recommended best practice is to identify the referenced resource by means of a label or number conforming to a formal identification system. | DescribedItemIRM nodes ./ddms:Resource/ddms:relatedResources MetaCardIRM nodes Not Applicable |
| Rights | Information about rights held in and over the resource. Typically, rights will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the rights element is absent, no assumptions may be made about any rights held in or over the resource. | DescribedItemIRM nodes ./ddms:Resource/ddms:rights MetaCardIRM nodes Not Applicable |

| Abstract Data Element | Definition | XPath and XML implementation notes |
|------------------------|--|--|
| Resource Security Mark | <p>The overall security classification and security handling instructions carried by the resource.</p> <p>Resource Security Mark applies to the resource-level classification, SCI controls, dissemination controls, non-IC markings, and other security provisions prescribed by Executive Order 13526, as amended, the Information Security Oversight Office (ISOO) Directive 1 of the National Archives and Records Administration, and the Intelligence Community marking standard maintained by the Controlled Access Program Coordination Office (CAPCO). These values are prominently presented, in the case of intelligence publications, at the top and bottom of every page and in other specified locations. See the <i>Intelligence Community Standard for Information Security Marking Metadata</i> for refinements of this conceptual element.</p> | <p>irm:ICResourceMetadataPackage/@ism:*/ntk:Access</p> <p>DescribedItemIRM nodes Not applicable</p> <p>MetaCardIRM nodes ./irm:NoticeList</p> |
| Source | <p>The resource from which the described resource is derived.</p> <p>The described resource may be derived from the related resource in whole or in part. Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system.</p> | <p>DescribedItemIRM nodes ./ddms:Resource/ddms:source</p> <p>MetaCardIRM nodes Not applicable</p> |
| Subject | <p>A topic of the resource.</p> <p>Typically, the topic will be represented using keywords, key phrases, or classification codes. Recommended best practice is to use a controlled vocabulary. To describe the spatial, temporal or virtual topic of the resource, use the Coverage element.</p> | <p>DescribedItemIRM nodes ./ddms:Resource/ddms:subjectCoverage ./ddms:Resource/ddms:subjectCoverage/ddms:Subject/ddms:keyword ./irm:ProductionMetricsList/irm:ProductionMetric ./irm:NonStateActorCoverageList/irm:NonStateActor</p> <p>MetaCardIRM nodes Not applicable</p> |
| Title | <p>A name given to the resource.</p> <p>Typically, a Title will be a name by which the resource is formally known.</p> | <p>DescribedItemIRM nodes ./ddms:Resource/ddms:title ./ddms:Resource /ddms:subtitle</p> <p>MetaCardIRM nodes Not applicable</p> |

| Abstract Data Element | Definition | XPath and XML implementation notes |
|--|--|---|
| Type | The nature or genre of the content of the resource. The Type includes terms describing general categories, functions, genres, or aggregation levels for content. Examples of Types include publication forms (e.g., reports or articles) and intelligence disciplines (e.g., SIGINT, MASINT, HUMINT). Recommended best practice is to use a controlled vocabulary. To describe the file format, physical medium, or dimensions of the resource, use the Format element. | DescribedItemIRM nodes ./ddms:Resource/ddms:type ./IntelType ./ProductLine MetaCardIRM nodes Not applicable |
| Records Management Information (Provisional) | Required information primarily supporting federal record keeping requirements. | DescribedItemIRM nodes ./irm:VitalRecordIndicator ./irm:RecordKeeper MetaCardIRM nodes ./irm:VitalRecordIndicator ./irm:RecordKeeper |

2.2 Additional guidance

This section provides additional guidance for encoding data in specific situations. The content of this section will evolve over time as additional situations are identified. Implementers of this DES are encouraged to contact the maintainers of this DES for further guidance when necessary.

2.2.1 DescribedItemIRM & MetaCardIRM

MetaCardIRM and **DescribedItemIRM** are two major components of a single IRM document and are not intended to be used as stand-alone documents. Each of them describes part of the IRM as a whole. From a Library Card analogy, the **ICResourceMetadataPackage** is the entirety of the "Library Card", the **DescribedItemIRM** contains information about the "book" while the **MetaCardIRM** contains information about the "Library Card"

There may be instances in which the author of the book documented in **DescribedItemIRM** and the author of the Library Card documented in **MetaCardIRM** are the same. In those cases the metadata may seem redundant. In the case where they are different, it becomes clear that an organization may create a

book documented in **DescribedItemIRM** while an entirely different agency may create the Library Card documented in **MetaCardIRM**.

MetaCardIRM has an **IdentifierList** easily confused with the **DescribedItemIRM/ddms:Resource /ddms:identifier**. These are similar constructs but for different purposes. Using the Library Card analogy again **ddms:identifier** inside **DescribedItemIRM** identifies the "book" maybe an International Standard Book Number (ISBN) number while the **IdentifierList** inside the **MetaCardIRM** identifies the "Library Card" with a unique identifier for the card. Since the **ICResourceMetadataPackage** may be in and of itself a classified document it needs its own identification for tracking, revision-recall, and auditing purposes.

2.2.2 DocumentID

For the purposes of the IC there needs to be a single document identifier that all documents will have. This document ID is denoted using the DDMS constructs by having a qualifier of "IC-ID" placed on a **ddms:identifier** element. The document identifier should be unique to this document across the whole of the IC. There is no central registry or managing body for document identifiers across the IC so it is the responsibility of individual producers to coordinate properly.

2.2.3 ISM Attribute usage

Both IRM and DDMS have adopted the recommended usage of the ISM resource attribute group being used on the root node of their schemas. Because of this decision, both the ISM attributes on the root node of IRM and those on the **ddms:resource** represent the classification attributes for all of their child elements that do not have **ism:excludeFromRollup='true'**. The only element in IRM that has the **excludeFromRollup='true'** is the Security element in DDMS. This is because the security element represents the classification information about the described item and not the classification of any content in the IRM.

2.2.4 Specification of publishing organization

The elements **ddms:publisher** and **irm:PublishingAgency** are used to identify the entity(ies) primarily responsible for releasing the information to the enterprise. The entity(ies) of interest in this context are foremost the organization responsible for the actual distribution of the data. The organizations and/or individuals responsible for creating the information are captured within the **ddms:creator** and **ddms:contributor** structures. The publishing organization's approved identifier value

is captured in an element called **irm:PublishingAgency**. Further decomposition of the **irm:PublishingAgency** is captured in the **irm:SubAgency** element. Depending on the enterprise requirement being addressed, a complete understanding of the Publisher requires evaluating the **irm:PublishingAgency/@irm:acronym** and **irm:SubAgency** value as well as the values found in the **ddms:affiliation** of the **ddms:publisher**, **ddms:creator** and **ddms:contributor** elements.

The **ddms:publisher** structure provides the ability to identify multiple levels of organizational structure and multiple organizations or individuals responsible for creating the information. The most basic ability to identify is captured with the required element **irm:PublishingAgency** using the attribute **acronym**. The controlled vocabulary enumeration (CVE) for **irm:acronym** includes values representing the organizations officially designated as part of the IC as defined in the DNI's *Overview of the United States Intelligence Community for the 111th Congress* of 2009, plus the DNI, plus additional entries intended to recognize non-IC publishers whose information is commonly used in support of the intelligence mission. One of these values must be selected.

In many cases, the AgencyAcronym CVE only includes the highest level of the organization structure (e.g., DNI), service or agency (e.g., US Army, DHS, DoS), or non-IC designation (e.g., OtherDoD, Foreign). In order to identify a Publisher at a level below what the AgencyAcronym CVE allows, use the **irm:SubAgency** element of the **irm:PublishingAgency**.

For consistency, populate **irm:SubAgency** with an approved organization acronym designator for the sub-organization. For multiple levels of sub-organization, list the acronyms in descending order delimited with the "/" character.

In cases where non-IC information (e.g., OtherDoD, OtherUSG, SLT, Foreign) is shared with the intelligence enterprise, the **irm:PublishingAgency/irm:acronym** should reflect the organization, which last prepared the information for consumption (e.g., converted the content into PUBS.XML, applied enhanced information resource metadata tagging, translated, or packaged the information into an official IC product) and shared the product with the enterprise. As that organization is affecting the record status of the product, it must take responsibility for addressing any questions about the information.

If a non-IC producer is providing information that is already compliant with IC enterprise data encoding standards, then the **irm:PublishingAgency/irm:acronym** should reflect the appropriate non-IC organization designator and the non-IC organizations office in the **irm:SubAgency** element. Examples of this scenario might exist in a USG department where there are sub-organizations designated in the IC and sub-organizations not in the IC; DoD where some sub-organizations support DIA, some

support a service, and some are not in the IC; State, Local, Tribal organizations with information that flows into the intelligence enterprise via DHS, NCTC, or other means; or with our foreign partners. In the case of foreign partners designations in the **irm:SubAgency**, precede the office acronym with the country code tri-graph in order to ensure uniqueness.

Examples

For NCTC:

```
<irm:PublishingAgency irm:acronym="DNI">  
<irm:SubAgency>NCTC</irm:SubAgency></irm:PublishingAgency>
```

For the XYZ component of NCTC:

```
<irm:PublishingAgency irm:acronym="DNI">  
<irm:SubAgency>NCTC/XYZ</irm:SubAgency></irm:PublishingAgency>
```

For the XYZ component of CIA:

```
<irm:PublishingAgencyirm:acronym="CIA">  
<irm:SubAgency>XYZ</irm:SubAgency></irm:PublishingAgency>
```

For the United State Postal Service:

```
<irm:PublishingAgency irm:acronym="OtherUSG">  
<irm:SubAgency>USPS</irm:SubAgency></irm:PublishingAgency>
```

For the JIOC at PACOM:

```
<irm:PublishingAgency irm:acronym="DIA">  
<irm:SubAgency>PACOM/JIOC</irm:SubAgency></irm:PublishingAgency>
```

For the J4 at PACOM:

```
<irm:PublishingAgency irm:acronym="OtherDoD">  
<irm:SubAgency>PACOM/J4</irm:SubAgency></irm:PublishingAgency>
```

2.2.5 MIME type

The Multipurpose Internet Mail Extensions (MIME) type for a IRM.XML document is application/dni-irm+xml. This is a convention for our community it has NOT been registered with the Internet Assigned Numbers Authority (IANA). Should there be a conflict in the future it will be addressed at that time. Systems can use this MIME type to facilitate communications and address business needs within the community.

Chapter 3 – XML Schema Guide

The detailed description and reference documentation for the IRM.XML schema can be found in a separate document entitled *IRM.XML Schema Guide*. This guide serves as an interactive presentation of the IRM.XML schema as well as a data element dictionary.

The guide was generated with a commercially available product named *oXygen®*, produced by SyncRO Soft.

The guide provides an interactive index to:

- Global Elements and Attributes
- Local Elements and Attributes
- Simple and Complex Types
- Groups and Attribute Groups
- Referenced Schemas

Where applicable, the guide provides:

- Diagram
- Namespace
- Type
- Children
- Used by
- Properties
- Patterns
- Enumerations
- Attributes
- Annotations
- Source Code

The guide is published in a folder consisting of a master HTML file with supporting graphics.

Chapter 4 – Constraint Rules

Constraint rules explicitly define the validation constraints for IRM.XML. They provide additional restrictions (i.e., constraints) on how the data should be structured and encoded, especially for criteria that exceed the constraints implemented in the XML Schema. These rules are written in plain English phrases; however, knowledge of the IRM.XML schemas is required to understand the rules. These constraint rules will eventually be offered in a more declarative form, such as Schematron. Complex constraint rules may be followed by text labeled **Human Readable**. This text is intended to inform the intent of the more formal language above it. Implementers are intended to implement the formal language, and should there be a perception of conflict, bring it to the attention of the appropriate configuration control body to be resolved.

4.1 Basics

The IRM.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.

4.1.1 Classified or Controlled Constraint Rules

Additional rules that are either classified or have handling controls can be found in separate annexes closely associated with the DES artifacts wherever they are located.

4.1.2 Terminology

For the purposes of this document, the following statements apply:

- The term “is specified” indicates that an attribute is applied to an element and the attribute has a non-null value.
- The term “must be specified” indicates that an attribute must be applied to an element and the attribute must have a non-null value.
- The term “is not specified” indicates that an attribute is not applied to an element, or an attribute is applied to an element and the attribute has a null value.

- The term “must not be specified” indicates that an attribute must not be applied to an element.

4.1.3 Rule Identifiers

Each constraint rule has an assigned rule ID, indicated in brackets preceding the constraint rule description. The rule IDs from 00001 to 10000 are unclassified and 10001 to 20000 are “for official use only” (FOUO). IDs from 20001 to 30000 are reserved for “Secret” rules and 30001 and above for more classified rules.

As the constraint rules are managed over time, ID’s from deleted rules will not be reused.

4.1.4 Errors and Warnings

The severity of a constraint rule violation is categorized as either an “Error” or a “Warning.” An “Error” is naturally more severe and is indicative of a clear violation of an IRM.XML constraint rule, which would be likely to have a significant impact on the quality of a document. A “Warning” is less severe although noteworthy, and may not necessarily have any impact on the quality of a document. The severity of a constraint rule violation is indicated in brackets preceding each constraint rule description

Each system responsible for processing a document (e.g., create, modify, transform, or exchange) must make a mission-appropriate decision about using a document with errors or warnings based on mission needs.

4.2 Global Constraints

[IRM-ID-00002][Error] For every optional attribute that is used in a document, a non-null value must be present.

Human Readable: In other words, the attribute “is specified” as defined in **Section 4.1.2.**

4.3 DES Constraints

The DES version is specified through attributes on the root element. The Schema constrains the values of these attributes. The DES version enables systems processing an instance document to be certain which set of schema, CVE’s and business rules are intended by the author to be used for any particular instance document.

4.4 Value Enumeration Constraints

[IRM-ID-00001][Error] If element **ddms:countryCode** has attribute **ddms:qualifier** specified as “urn:us:gov:ic:cvenum:irm:coverage:fips:digraph:v1” then the value of attribute **ddms:value** must be in CVEnumIRMCoverageFIPSDigraph.xml.

Human Readable: FIPS CountryCodes must be in the FIPS CVE

[IRM-ID-00002][Error] If element **ddms:countryCode** has attribute **ddms:qualifier** specified as “urn:us:gov:ic:cvenum:irm:coverage:iso3166:digraph:v1” then the value of attribute **ddms:value** must be in CVEnumIRMCoverageISO3166Digraph.xml.

Human Readable: ISO digraph CountryCodes must be in the ISO digraph CVE

[IRM-ID-00003][Error] If element **ddms:countryCode** has attribute **ddms:qualifier** specified as “urn:us:gov:ic:cvenum:irm:coverage:iso3166:trigraph:v1” then the value of attribute **ddms:value** must be in CVEnumIRMCoverageISO3166Trigraph.xml.

Human Readable: ISO trigraph CountryCodes must be in the ISO trigraph CVE

[IRM-ID-00004][Error] If element **ddms:countryCode** has attribute **ddms:qualifier** specified as “urn:us:gov:ic:cvenum:irm:coverage:iso3166:numeric:v1” then the value of attribute **ddms:value** must be in CVEnumIRMCoverageISO3166Numeric.xml.

Human Readable: ISO numeric CountryCodes must be in the ISO numeric CVE

[IRM-ID-00005][Error] If element **ddms:language** has the attribute **ddms:qualifier** value of “urn:us:gov:ic:cvenum:irm:iso639:digraph:v1” then the value of attribute **ddms:value** must be in CVEnumIRMISO639Digraph.xml and no country code portion may be specified in the **Language** element value.

Human Readable: ISO 639 digraph language codes must be in the ISO 639 digraph CVE

[IRM-ID-00006][Error] If element **ddms:language** has the attribute **ddms:qualifier** value of “urn:us:gov:ic:cvenum:irm:iso639-2:trigraph:v1” then the value of attribute **ddms:value** must be in CVEnumIRMISO639-2Trigraph.xml and no country code portion may be specified in the **ddms:value** value.

Human Readable: ISO 639-2 trigraph language codes must be in the ISO 639-2 trigraph CVE

[IRM-ID-00007][Error] If element **ddms:language** has the attribute **ddms:qualifier** value of “urn:us:gov:ic:cvenum:irm:iso639-3:trigraph:v1” then the value of attribute **ddms:value** must be in CVEnumIRMISO639-3Trigraph.xml and no country code portion may be specified in the **ddms:value** value.

Human Readable: ISO 639-3 trigraph language codes must be in the ISO 639-3 trigraph CVE

[IRM-ID-00008][Error] If element **ddms:language** has the attribute **ddms:qualifier** value of “RFC1766” then the language code portion of the **ddms:value** attribute value must be in CVEnumIRMISO639Digraph.xml and the country code portion, if present, must be in CVEnumIRMCoverageISO3166Digraph.xml.

Human Readable: RFC1766 language codes must comply with the RFC by using parts from ISO 639 Digraph and ISO 3166 Digraph

[IRM-ID-00009][Error] If element **ddms:language** has the attribute **ddms:qualifier** value of “RFC3066” then the language code portion of the **ddms:value** attribute value must be in CVEnumIRMISO639Digraph.xml or CVEnumIRMISO639-2Trigraph.xml and the country code portion, if present, must be in CVEnumIRMCoverageISO3166Digraph.xml.

Human Readable: RFC3066 language codes must comply with the RFC by using parts from ISO 639 Digraph or ISO 639-2 Trigraph and ISO 3166 Digraph

[IRM-ID-00010][Error] If element **ddms:language** has the attribute **ddms:qualifier** value of “RFC4646” then the language code portion of the **ddms:value** attribute value must be in CVEnumISO639Digraph.xml or CVEnumISO639-2Trigraph.xml and the country code portion, if present, must be in CVEnumIRMCoverageISO3166Digraph.xml.

Human Readable: RFC4646 language codes must comply with the RFC by using parts from ISO 639 Digraph or ISO 639-2 Trigraph and ISO 3166 Digraph

4.5 General Constraints

[IRM-ID-00011][Error] If an element **irm:MetaCardIRM/irm:IdentifierList/ddms:identifier** does not exist with the attribute **ddms:qualifier** value of “IC-ID”.

Human Readable: Every metacard must have an ID identified as an IC-ID

[IRM-ID-00012][Error] If more than 1 element

irm:MetaCardIRM/irm:IdentifierList/ddms:identifier exists with the attribute **ddms:qualifier** value of “IC-ID”.

Human Readable: Every metacard must have only 1 ID identified as an IC-ID

[IRM-ID-00013][Error] If an element **irm:DescribedItemIRM/ddms:Resource/ddms:identifier** does not exist with the attribute **ddms:qualifier** value of “IC-ID”.

Human Readable: Every DescribedItemIRM must have an ID identified as an IC-ID

[IRM-ID-00014][Error] If more than 1 element

irm:DescribedItemIRM/ddms:Resource/ddms:identifier exists with the attribute **ddms:qualifier** value of “IC-ID”.

Human Readable: Every DescribedItemIRM must have only 1 ID identified as an IC-ID

[IRM-ID-00015][Error] If element **ddms:dates** exists without one of the attributes **ddms:created** or **ddms:posted**

Human Readable: Every date must have 1 of created or posted.

[IRM-ID-00016][Error] The permissible values for the year range are 1901 through the current year for attributes **irm:dateApproved**, **ddms:infoCutOff**, **ddms:posted**, and **ddms:created**

Human Readable: Dates must be after 1901 and in the past for **irm:dateApproved**, **ddms:infoCutOff**, **ddms:posted**, and **ddms:created**.

[IRM-ID-00017][Error] The permissible values for the year range are 1901 through 9999 for element **ddms:validTil**.

Human Readable: **ddms:validTil** must be after 1901.

[IRM-ID-00018][Error] The permissible values for the decimal seconds are .0 through .999 for attributes **irm:dateApproved**, **ddms:start**, **ddms:end**, **ddms:infoCutOff**, **ddms:posted**, **ddms:validTil** and **ddms:created**.

[IRM-ID-00019][Warning] **irm:dateApproved** must not be later than **ddms:created**, and **ddms:posted**.

[IRM-ID-00020][Error] **ddms:infoCutOff** must not be later than **ddms:posted**, and **ddms:created**.

[IRM-ID-00021][Warning] **ddms:validTil** must not be earlier than **irm:dateApproved**, **ddms:infoCutOff**, **ddms:posted**, and **ddms:created**.

[IRM-ID-00022][Error] For any element **TimePeriod** **start** must not be later than **end**.

[IRM-ID-00023][Error] The permissible values for the year range are 0001 through 9999 for elements **ddms:start** and **ddms:end**.

Human Readable: **ddms:start** and **ddms:end** must be positive integers less than 10,000.

[IRM-ID-00024][Warning] For elements **irm:dateApproved**, **ddms:start**, **ddms:end**, **ddms:infoCutOff**, **ddms:posted**, **ddms:validTil** and **ddms:created**., if the time designator (T) is specified, it is recommended that time zone be specified.

4.6 Information Security Markings (ISM.XML)

Most constraint rules specific to the application of information security markings are documented in the *XML Data Encoding Specification for Information Security Marking*

Metadata and related documents. The rules in this section are additional constraints on the specific implementation of ISM in IRM.XML

[IRM-ID-00025][Error] The attribute **ism:excludeFromRollup** must not be specified for any element in the namespace <http://metadata.dod.mil/mdr/ns/DDMS/3.0/> except **security**.

[IRM-ID-00026][Error] The root element must have the attribute **ism:compliesWith** in the namespace urn:us:gov:ic:ism containing the value of "ICD-710".

Human Readable: All IRM.XML documents must comply with the applicable ICD-710 rules encoded in ISM.

Appendix A 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 Intelligence Community Chief Information Officer (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 Intelligence Community Standard (ICS) 500-20.

Appendix B Acronyms

ADD - Abstract Data Definition

CAPCO – Controlled Access Program Coordination Office

CVE – Controlled Vocabulary Enumeration

DCMI – Dublin Core Metadata Initiative

DC MES – Dublin Core Metadata Element Set

DES – Data Encoding Specification

DOI – Digital Object Identifier

FOUO - For Official Use Only

GNS – Geographic Names Server

HTML – HyperText Markup Language

IC CIO – Intelligence Community Chief Information Officer

ICD – Intelligence Community Directive

ICS – Intelligence Community Standard

IRM.XML - XML Data Encoding Specification for Information Resource Metadata

ISBN – International Standard Book Number

ISO – International Organization for Standardization

ISOO – Information Security Oversight Office

KA – Knowledge Assertion

KOS – Knowledge Organization System

MIME – Internet Media Types

NARA – National Archives and Records Administration

NGA – National Geospatial Intelligence Agency

NTK - Need To Know

OCIO - Office of the Intelligence Community Chief Information Officer

ODNI – Office of the Director of National Intelligence

TGN – Thesaurus of Geographic Names

URI – Uniform Resource Identifier

URL – Uniform Resource Locator

W3CDTF – World Wide Web Consortium Date Time Format

XML – Extensible Markup Language

Appendix C Glossary

No pertinent glossary items requiring further definition.

Appendix D Bibliography

This appendix lists all the sources referenced in this DES and lists other sources that may have been used in other DES. This appendix is a shared resource across multiple documents so in any given DES there are likely sources that are not referenced in that particular DES.

(CAPCO Implementation Guide)

Intelligence Community Classification and Control Markings Implementation Manual. Unclassified FOUO version. Volume 3, Edition 1 (Version 3.1). 7 May 2010. Director of National Intelligence (DNI), Special Security Center (SSC), Controlled Access Program Coordination Office (CAPCO).

https://www.intelink.gov/sites/ssc/divisions/capco/CAPCO%20Resources/CAPCO_Implementation%20Manual_FOUO_v3%201_07%20May%2010%20.pdf

(CAPCO Register)

Authorized Classification and Control Markings Register. Unclassified FOUO version. Volume 3, Edition 1 (Version 3.1). 7 May 2010. Director of National Intelligence (DNI), Special Security Center (SSC), Controlled Access Program Coordination Office (CAPCO).

[https://www.intelink.gov/sites/ssc/divisions/capco/CAPCO%20Resources/CAPCO_Register_FOUO_v3%201_07%20May%2010%20\(2\).pdf](https://www.intelink.gov/sites/ssc/divisions/capco/CAPCO%20Resources/CAPCO_Register_FOUO_v3%201_07%20May%2010%20(2).pdf)

(DC MES)

Dublin Core Metadata Element Set. Version 1.1. 02 June 2003. Dublin Core Metadata Initiative. <http://dublincore.org/documents/dces/>.

(E.O. 12958, as amended)

Executive Order 12958 – Classified National Security Information, as Amended. Federal Register, Vol. 68, No. 60. 25 March 2003. The White House.

<http://www.archives.gov/isoo/policy-documents/eo-12958-amendment.html>.

(E.O. 12829, as amended)

Executive Order 12829 – National Industrial Security Program, as Amended. Federal Register, Vol. 58, No. 240. 16 December 1993. The White House.

<http://www.archives.gov/isoo/policy-documents/eo-12829.html>

(E.O. 13526)

Executive Order 13526 – Classified National Security Information 29 December 2009. The White House. <http://www.archives.gov/isoo/pdf/cnsi-eo.pdf>.

(ICD 206)

Sourcing Requirements for Disseminated Intelligence Products. Intelligence Community Directive Number 206. 17 October 2007. Office of the Director of National Intelligence. http://www.dni.gov/electronic_reading_room/ICD_206.pdf.

(ICD 500)

Director of National Intelligence Chief Information Officer. Intelligence Community Directive Number 500. 7 August 2008. Office of the Director of National Intelligence. http://www.dni.gov/electronic_reading_room/ICD_500.pdf.

(ICD 710)

Classification and Control Markings System. Intelligence Community Directive Number 710. 11 September 2009. Office of the Director of National Intelligence. http://www.dni.gov/electronic_reading_room/ICD_710.pdf

(ISO 639-2)

Codes for the representation of names of languages – Part 2: Alpha-3 code. ISO 639-2:1998. International Organization for Standardization (ISO). http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=4767.

(ISO 3166-1)

Codes for the representation of names of countries and their subdivisions – Part 1: Country codes. ISO 3166-1:2006. International Organization for Standardization (ISO). http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39719.

(ISO 8601)

Data elements and interchange formats – Information interchange – Representation of dates and times. ISO 8601:2004. International Organization for Standardization (ISO). http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40874.

(ISO 15836)

Information and documentation – The Dublin Core metadata element set. ISO 15836:2009. International Organization for Standardization (ISO). http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=52142.

(ISOO Directive 1)

Classified National Security Information (Directive No. 1); Final Rule. 32 CFR Parts 2001 and 2004. Federal Register, Vol. 68, No. 183. 22 September 2003. Information Security Oversight Office (ISOO), National Archives and Records Administration (NARA). <http://www.archives.gov/isoo/policy-documents/eo-12958-implementing-directive.pdf>.

(RFC 3066)

Tags for the Identification of Languages. January 2001. H. Alvestrand. Cisco Systems. <http://www.rfc-editor.org/rfc/rfc3066.txt>.

Appendix E Points of Contact

This technical specification is managed by the Office of the Intelligence Community Chief Information Officer (IC CIO). As of this writing, the IC CIO/IC Enterprise Architecture (ICEA) Directorate facilitates the IC data collaboration and coordination forums responsible for the selection or development of common IC technical data specifications. Direct all inquiries about this IC technical specification to IC CIO/ICEA, the IC's data collaboration and coordination forum, or IC element representatives involved in those forums.

Appendix F Change History

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

Table 2. DES Version Identifier History

| Version | Identifier | Date | Purpose |
|---------|-----------------------|------------------|---|
| 1.0 | IRM V1.0 | July 2009 | Initial Release |
| 2 | ICTechSpec 500.D.5-V2 | 7 September 2010 | Routine revision to technical specification. For details of changes, see appendix F.2 |
| 3 | ICTechSpec 500.D.5-V3 | 6 December 2010 | Routine revision to technical specification. For details of changes, see appendix F.1 |

F.1 V3 Change Summary

The following table summarizes the changes made to V2 in developing V3

Table 3. Data Encoding Specification V3 Change Summary

| Change | Artifacts changed | Compatibility Notes |
|--|-------------------------|---|
| Use ISM V5 | Schema | Data generation and Ingestion systems need to be updated to properly enforce the new constraint rule |
| Add IRM.XML MIME type | DES, Schema | IRM.XML MIME type has been declared in order to facilitate communications and address business needs within the community |
| Remove Appendix H Reading the Schematics | Documentation | Knowledge of how to interpret these schema images is common making this appendix unnecessary. |
| Add support for expressing coverage of NonState Actors | Documentation Schema | Data generation and Ingestion systems need to be updated to properly support new elements. |

F.2 V2 Change Summary

The following table summarizes the changes made to V1 in developing V2

Table 4. Data Encoding Specification V2 Change Summary

| Change | Artifacts changed | Compatibility Notes |
|--|-------------------|---|
| Added all constructs other than ddms:Resource | All | Prior data will need to have the constructs other than ddms:Resource and will have to map ddms:Resource to irm:ICResourceMetadataPackage |

Appendix G Configuration Management

The selection or development of technical data specifications of common interest to the IC are collaborated and coordinated currently within governance forums managed by the IC CIO. Change requests for this technical data specification should be directed to the office identified in **Appendix E – Points of Contact**.

Appendix H Controlled Vocabulary Enumerations

The Controlled Vocabulary Enumerations (CVEs) used in this DES are as follows:

| CVE File name | Definition | Attribute and Rules Cross reference |
|---------------------------------------|---|---|
| CVEnumIRMCoverageFIPSDigraph | All currently valid FIPS digraphs. | CountryCode PUBS-ID-00066 PUBS-ID-00011 |
| CVEnumIRMCoverageISO3166Digraph | All currently valid ISO-3166 digraphs. | CountryCode PUBS-ID-00067 PUBS-ID-00059 PUBS-ID-00060 PUBS-ID-00061 |
| CVEnumIRMCoverageISO3166Numeric | All currently valid ISO-3166 numeric values | CountryCode PUBS-ID-00069 |
| CVEnumIRMCoverageISO3166Trigraph | WWW and All currently valid ISO-3166 trigraphs | CountryCode PUBS-ID-00068 |
| CVEnumIRMCoverageISO4217Numeric | All currently valid ISO-4217 Numeric codes. | @unitOfMeasure PUBS-ID-00064 |
| CVEnumIRMCoverageISO4217Trigraph | All currently valid ISO-4217 Trigraphs. | @unitOfMeasure PUBS-ID-00063 |
| CVEnumIRMCoverageISO639Digraph | All currently valid ISO-639-1 Digraphs. | Language PUBS-ID-00056 PUBS-ID-00059 PUBS-ID-00061 |
| CVEnumIRMCoverageISO639-2Trigraph | All currently valid ISO-639-2 Trigraphs. | Language PUBS-ID-00057 PUBS-ID-00060 |
| CVEnumIRMCoverageISO639-3Trigraph | All currently valid ISO-639-3 Trigraphs. | Language PUBS-ID-00058 |
| CVEnumIRMAgencyAcronym | All currently valid Agency Acronyms for use with publisher. | AgencyAcronym PUBS-ID-00076 |
| CVEnumMimeType | All currently valid MIME Types. | @MIMEtype PUBS-ID-00074 PUBS-ID-00075 |
| CVEnumIRMIntelSubDisciplineTechniques | All currently valid Intel sub Discipline technique codes. | IntelSubdisciplineTechnique SchemaEnumeration |
| CVEnumIRMIntelSubDisciplines | All currently valid Intel sub Discipline codes. | IntelSubdiscipline SchemaEnumeration |
| CVEnumIRMIntelDisciplines | All currently valid Intel Discipline codes. | IntelDiscipline SchemaEnumeration |
| CVEnumIRMProductionMetricsCoverage | All currently valid Production Metrics Coverage | ProductionMetricCoverage SchemaEnumeration |
| CVEnumIRMProductionMetricsSubject | All currently valid Production Metrics Subjects | ProductionMetricSubject SchemaEnumeration |