

XML DATA ENCODING SPECIFICATION FOR INTELLIGENCE PUBLICATIONS VERSION 6 (PUBS.XML.V6)

ICTechSpec 500.D.3-V6

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

11 April 2011



Table of Contents

Table of Contents	i
List of Tables	ii
Chapter 1 – Introduction	1
1.1 Purpose.....	1
1.2 Needs and Requirements.....	1
1.3 Audience and Applicability	2
1.4 Utility	2
1.5 Version Information	2
1.6 Components of this Data Encoding Specification	2
1.7 Normative and Informative Components.....	3
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
2.1.1 <i>Publication Metadata</i>	5
2.1.2 <i>Information Resource Metadata</i>	9
2.1.3 <i>Source Reference Citation Metadata</i>	13
2.1.4 <i>Information Security Marking Metadata</i>	17
2.2 Additional guidance.....	18
2.2.1 <i>Specification of publishing organization</i>	18
2.2.2 <i>Specification of Need-To-Know Access Parameters</i>	20
2.2.3 <i>Specification of Production Metrics Reporting Metadata</i>	20
2.2.4 <i>Specification of PublicationMetadataList</i>	20
2.2.5 <i>Specification of ism:notice</i>	21
2.2.6 <i>Specification of approximable dates</i>	22
2.2.7 <i>Approximable Dates in constraint rules</i>	23
2.2.8 <i>MIME type</i>	23
Chapter 3 – XML Schema Guide.....	24
Chapter 4 – Data Validation Constraint Rules.....	25
4.1 Basics.....	25
4.1.1 <i>Schematron</i>	25
4.1.2 <i>Classified or Controlled Constraint Rules</i>	26
4.1.3 <i>Terminology</i>	26
4.1.4 <i>Rule Identifiers</i>	26
4.1.5 <i>Errors and Warnings</i>	27
4.2 Non-null Constraints.....	27
4.3 DES Constraints	27
4.4 Inherited Constraints.....	27
4.5 Value Enumeration Constraints	28
4.6 Dates and Times.....	28
4.7 Time Zone Indicators	30
4.8 Information Security Markings (ISM.XML).....	30
4.9 Obsolete rule numbers	30

Chapter 5 – Data Rendering Constraint Rules	32
5.1 Rule Identifiers	32
5.2 Errors and Warnings	32
Appendix A IC CIO Approval Memo.....	34
Appendix B Acronyms	35
Appendix C Glossary	37
Appendix D Bibliography.....	38
Appendix E Points of Contact.....	41
Appendix F Change History.....	42
F.1 V6 Change Summary.....	43
F.2 V5 Change Summary.....	45
F.3 V4 Change Summary.....	46
F.4 V3 Change Summary.....	48
F.5 V2 Change Summary.....	49
Appendix G Configuration Management.....	51
Appendix H Controlled Vocabulary Enumerations	52
Appendix I IRM.HTML to PUBS.XML Mappings	54

List of Tables

Table 1. Publication Structure Abstract Data Elements to Physical XML Elements	6
Table 2. Section Structure Abstract Data Elements to Physical XML Elements.....	6
Table 3. Narrative Structure Abstract Data Elements to Physical XML Elements	7
Table 4. Information Resource Metadata Conceptual Elements to Physical XML Elements	9
Table 5. Source Reference Citation Abstract Data Elements to Physical XML Elements.	13
Table 6. Source Reference Citation Abstract Data Element Refinements to Physical XML Elements	14
Table 7. Date/Time-Related Data Types and Layout Representations	29
Table 8. DES Version Identifier History	42
Table 9. Data Encoding Specification V6 Change Summary	43
Table 10. Data Encoding Specification V5 Change Summary	45
Table 11. Data Encoding Specification V4 Change Summary	46
Table 12. Data Encoding Specification V3 Change Summary	48
Table 13. Data Encoding Specification V2 Change Summary	49
Table 14. CVE Definitions	52
Table 15. Occurrence codes used in mapping table	55
Table 16. IRM.HTML to PUBS.XML XPath Mappings.....	55

Chapter 1 – Introduction

1.1 Purpose

This *XML Data Encoding Specification for Intelligence Publications* (PUBS.XML) defines detailed specifications for using Extensible Markup Language (XML) to encode publication metadata in compliance with the IC Abstract Data Definition (IC.ADD). This Data Encoding Specification (DES) defines the XML elements and attributes, associated structures and relationships, mandatory and cardinality requirements, and permissible values for representing the publications concepts using XML. It is the umbrella XML standard for text-based intelligence products.

1.2 Needs and Requirements

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

- Improving publication and dissemination efficiency by reducing the cost and time for performing manual and complex rendering, manipulation, and content transformation of information resource metadata in context of an intelligence publication.
- Facilitating discovery and exchange of content consisting of mostly text supplemented by interspersed non-textual content (i.e., multi-media) information between collectors, all-source analysts, and consumers.
- Capturing an intelligence publication's overall security marking metadata in order to support attribute and clearance-based information management practices, such as: secure collaboration; content management; content and portion-level filtering of discovery results, and content transfers across security domains.
- Capturing source reference citations to provide intelligence collectors the ability to systematically analyze how and how often the data they gather or produce is being used in order to facilitate better management of collection and production resources.
- Capturing source reference citations to enhance the analytic integrity of formally disseminated intelligence information and improving the traceability of collected information to analytic judgments and conclusions.
- Capturing and retaining a greater understanding of an intelligence publication's meaning, purpose, genesis, and characteristics as identified by a human or service.

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 data elements in the IC.ADD in a particular physical encoding (e.g., data or file format). For example:

- DESs for textual markup formats, such as Extensible Markup Language (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, for e.g. Microsoft Word, define document properties; styles; fields; cardinalities; processing requirements; and use.

1.5 Version Information

This is **Version 6** of this DES. This version number must be specified in the **DESVersion** attribute within any XML instance document claiming to be valid against this version. A separate **DESVersion** attribute must be specified for each DES against which an instance document is claiming compliance. These attributes must 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 Data Encoding Specification

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 PUBS.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 PUBS.XML beyond those in the XML Schema.
 - **Chapter 5 – Data Rendering Constraint Rules.** The constraint rules in this chapter define constraints of the rendering of PUBS.XML documents. The intent of this chapter is to inform the development of systems capable of rendering or displaying PUBS.XML data for use by individuals not familiar with the details of the PUBS.XML markup.

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.

Example files are also provided to illustrate various encoding situations. The revision/recall and source reference citation examples illustrate the situations presented in the former “Best Practices”.

1.7 Normative and Informative Components

The XML schemas, CVE values from the XML CVE files, and the Chapter 4 and 5 constraint rules are normative for this DES. The rest of this document, the descriptive content referenced within the XML Schema Guide, HTML CVE value files, and example files are informative.

1.8 Technical Encoding Dependencies

This DES relies on:

- Information Security Marking (ISM.XML.V6)
- Information Resource Metadata (IRM.XML.V4)
- Need To Know (NTK.XML.V4)
- 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.

2.1.1 Publication Metadata

1. Publication Structures (**Table 1**)
2. Section Structures (**Table 2**)
3. Narrative Structures (**Table 3**)

Table 1. Publication Structure Abstract Data Elements to Physical XML Elements

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Article	Used for news or journalist reporting and for other publications with little or no front or rear matter.	/IntelDoc
Report	Used for publications with extensive front and rear matter, and body matter that is subdivided into parts, chapters, and/or sections.	/IntelDoc

Table 2. Section Structure Abstract Data Elements to Physical XML Elements

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Section	Generic subdivision of an article, report, or correspondence.	//Section
Section Title	Generic section's primary title.	//Section/Title
Sidebar	A short article that is substantially parallel to the text of the main report but not directly a part of it.	//Sidebar
Appendix	A collection of supplementary material usually placed after the main body of writing.	//Appendix
Attachments	A section that is appended or attached to a main document, usually a correspondence document.	//Attachments
Collect Source	Information about sources from which intelligence is collected	//CollectSourceList
Bibliography	A list of the works referenced in the body of a publication or consulted by the author in its production.	//Bibliography //Bibliography/BibliographyDivision //Bibliography/BibliographyEntry
Distribution List	A series of addresses or routing symbols for distribution of a publication.	//DistributionList //DistributionEntry
Glossary	A list of often difficult or specialized words with their definitions.	//Glossary //Glossary/GlossaryDivision //Glossary/GlossaryEntry
Index	Index is an alphabetized list of names, places, and/or subjects that facilitates reference to the body of the publication.	//Index //Index/IndexDivision //Index/IndexEntry
Key Findings	Key conclusions reached after examination or investigation.	//KeyFindings

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Preface	A Preface is an introductory section offering information about the source of the request for a report, who wrote the report, the source of the information, how the study was conducted, etc. This element does not specifically address the scope of the report.	//Preface
Scope	The Scope is the extent or range of application, aim or purpose of a report.	//Scope
Summary	A comprehensive and usually brief abstract, recapitulation, or compendium of facts, statements, and/or findings.	//Summary
Table of Contents	Listing of sections, figures, tables, or other specially titled content listed by title within the publication and pointer to the content.	//TOC

Table 3. Narrative Structure Abstract Data Elements to Physical XML Elements

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Equation	A complex structure representing a formula or an expression, such as a mathematical or chemical equation.	//Equation
List	Series of items representing distinct but related thoughts written together in a meaningful grouping or sequence.	//List //List/ListItem
Media Resource	Media Resource is a complex structure including a form of presentable media (e.g., graphic, animation, video) and some form of unique identification (e.g., title) or clarification (e.g., legend).	//MediaResource
Note	Comment or explanation further clarifying surrounding content.	//Note //NoteInline //irm:NoticeList
Paragraph	A distinct portion of written matter dealing with a particular idea.	//Para
Quote	Passage copied verbatim from a book, speech or other source that is properly referenced.	//Quote
Table	Complex structure including a two-dimensional list organized into a grid containing rows and columns with special presentation characteristics and some form of unique identification (e.g., title).	//Table

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Assertion	A complex structure used to highlight content and associate special emphasis (via formatting), semantic understanding (via tagging) or ancillary value-added information (via hyperlink).	//Account //BEnumber //CommData //Commodity //Concept //CountryName //Date //DateTime //Drug //EntityUntyped //Equipment //Event //Facility //GeoFeature //GeoRef //Identifier //InfoBearer //LocationOfInterest //MilitaryUnit //Money //Nomenclature //Organization //Person //QuantityReference //SystemClass //Term //Time //Vehicle //Weapon
Footnote	A note that comments on—or cites a reference for—a designated part of the content, usually presented inline, at the bottom of the page, or at the end of a publication.	//Footnote
Source Citation	Bibliographic citation specialized to identify information sources necessary to substantiate analysis.	//SourceGroup //SourceGroup/SourceReference

2.1.2 Information Resource Metadata

All XPath expressions in this table, which fall under the PublicationMetadata element, are prefixed with the corresponding parent element of either **AdministrativeMetadata** or **DescriptiveMetadata**. XPath expressions without either of these prefixes, such as **/SourceGroup**, may appear in other contexts in a publication XML file.

Table 4. Information Resource Metadata Conceptual Elements to Physical XML Elements

Abstract Data Element	Abstract Data 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.	//AdministrativeMetadata/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 GEOnet 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.	//DescriptiveMetadata/Coverage //irm:ProductionMetricsList
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.	//AdministrativeMetadata/Publisher

Abstract Data Element	Abstract Data 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.	//AdministrativeMetadata/DateList
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.	//DescriptiveMetadata/Description
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.	//AdministrativeMetadata/RecordsManagementInfo/ApplicationSoftware //AdministrativeMetadata/RecordsManagementInfo/MediaFormat
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).	//AdministrativeMetadata/IdentifierList
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."	//DescriptiveMetadata/Language

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
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.	//AdministrativeMetadata/Publisher
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.	//AdministrativeMetadata/Relation
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.	//AdministrativeMetadata/Rights //MediaResource/CopyrightAttribution //ntk:Access
Resource Security Mark	<p>The overall security classification and security handling instructions carried by the resource.</p> <p>These values are prominently presented, in the case of 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>	/IntelDoc
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>	//SourceGroup
Subject	A topic of the resource. 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.	//DescriptiveMetadata/Activity //DescriptiveMetadata/Subject //irm:ProductionMetricsList

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Title	A name given to the resource. Typically, a Title will be a name by which the resource is formally known.	//DescriptiveMetadata/AbbreviatedTitle //DescriptiveMetadata/Title //DescriptiveMetadata/Subtitle
Type	<p>The nature or genre of the content of the resource.</p> <p>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.</p>	//DescriptiveMetadata/IntelType //DescriptiveMetadata/Note //OtherProperty //DescriptiveMetadata/PositiveIntel //DescriptiveMetadata/ProductLine //irm:NoticeList/irm:Notice[@ism:notice]] //NoteInline[@ism:notice]
Records Management Information (Provisional)	Required information primarily supporting federal record keeping requirements.	///AdministrativeMetadata/RecordsManagementInfo/VitalRecordIndicator //AdministrativeMetadata/RecordsManagementInfo/irm:RecordKeeper

2.1.3 Source Reference Citation Metadata

The conceptual elements defined in the ICS are mapped to XML elements as shown in **Table 5**. Conceptual element refinements are mapped to XML elements as shown in **Table 6**. In the mapping classification information is usually mapped to the Information Security Markup ISM and denoted as **@ism:*** to represent the collection of attributes defined by ISM.

Table 5. Source Reference Citation Abstract Data Elements to Physical XML Elements

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Bibliographic Resource (DCMI)	<p>A book, article, or other documentary resource.</p> <p>IC Expansion: In the context of source reference citations, a bibliographic resource is all significant and substantive reporting or other information upon which analytic judgment, assessments, estimates, or confidence levels depend. An intelligence product may be derived from one or more source references in whole or in part. Recommended best practice is to identify a related information resource by means of a formal identification system.</p>	An actual document being cited.
Bibliographic Citation (DCMI)	<p>A bibliographic reference for the [cited] resource.</p> <p>IC Expansion: A special type of bibliographic reference (i.e., a formal identification system) unique to the intelligence discipline that contains pertinent information resource metadata and details of the extent of the information being referenced. In accordance with ICD 206, source reference citations are to be listed in a special section at the end of intelligence products.</p>	<p>//SourceReference</p> <p>Except as otherwise specified, all XPath specifications in Table 6 are relative to //SourceReference.</p>

Table 6. Source Reference Citation Abstract Data Element Refinements to Physical XML Elements

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Citation Security Mark	Classification marking used for the overall <u>Bibliographic Citation</u> . This is the citation's portion mark as displayed in the bibliography or collection of source references.	@ism:*
Consulted	A date and time when a cited resource was used as a basis for analytic judgment.	DateTimeReferenced
Creator (DCMI)	An entity primarily responsible for making the [cited] resource. IC Expansion: The <u>Creator</u> can represent an author and/or coauthor and/or point of contact for the cited resource. The entity must be from or associated with the originating organization defined by the <u>Publisher</u> . If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the <u>Citation Security Mark</u> should be followed.	POCinfo – IC originator AuthorInfo – Human author(s) EditorInfo – Editor(s) Classification information is provided in: @ism:* within the appropriate element.
Date of Information	A date, time range, or time period representing the relative currency of the specific information cited.	DateInformation
Identifier (DCMI)	An unambiguous reference to the [cited] resource within a given context. IC Expansion: Recommended best practice is to identify a related information resource by means of a formal identification system. Examples might include a report serial number, document name or number, image frame identification code, or an organization internal identification or tracking number.	DocumentID – Unique resource ID SourceID – Agency internal identifier or tracking number
Issued (DCMI)	Date of formal issuance (e.g., publication) of the [cited] resource.	DatePublished
Link	A hyperlink to the cited resource. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the <u>Citation Security Mark</u> should be followed.	Link/@xlink:href or SegmentReferenced/Link/@xlink:href Classification information is provided in: Link/@ism:*

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Publisher (DCMI)	<p>An entity responsible for making the [cited] resource available.</p> <p>IC Expansion: An IC element, national government, international organization, or open-source owner(s) and/or producer(s) of a cited resource. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the <u>Citation Security Mark</u> should be followed.</p>	<p>IC originator: AgencyAcronym Open source: POCinfo Foreign government: CountryCode and optionally POCinfo</p> <p>The type of publisher (IC, open-source, or foreign), is provided in: @publisherType</p> <p>Classification information is provided in: AgencyAcronym/@ism:* or CountryCode/@ism:* or POCinfo/@ism:*</p>
Segment Referenced	<p>An identifier or description of the extent of the cited resource.</p> <p>Typically includes a form of label (e.g., a section or paragraph number, image feature, page number or range, video frame or range, etc.), possibly the classification of the extent, and possibly a link into the cited resource. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the <u>Citation Security Mark</u> should be followed.</p>	<p>SegmentReferenced/MediaExtent or SegmentReferenced/SegmentLabel or SegmentReferenced/Link/@xlink:href</p> <p>Classification information is provided in: SegmentReferenced/ReferencedResourceSecurity/@ism:*</p> <p>The security attributes should represent the contemporary CAPCO marks for the segment referenced. The security attributes on the ReferencedResourceSecurity element do NOT impact the overall security mark of the citation since they only convey what the classification of the segment referenced is not any data in the current document.</p> <p>Additionally if the segment referenced has a legacy CAPCO or other marking the text form of that marking could be put in SegmentReferenced/ReferencedResourceSecurity/OriginalClassificationMarking</p> <p>Should the OriginalClassification text be in and of itself classified that Classification information may be provided in: SegmentReferenced/ReferencedResourceSecurity/OriginalClassificationMarking/@ism:*</p>

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Sourced Content	A word, phrase, sentence, or other contiguous text string for which attribution is being cited. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the <u>Citation Security Mark</u> should be followed.	<p>//SourceGroup/SourcedText</p> <p>Classification information is provided in: //SourceGroup/SourcedText/@ism:*</p>
Source Descriptor	An explanation of factors contained in the cited resource or publicly available information that the producing organization assesses may affect the quality or reliability of the information in the specific cited resource. Factors may include, but are not necessarily limited to, completeness, precision or technical quality, context, or age/currency of the information. In the case of human sources, this explanation may include information that describes the level of access, past reporting record, or potential biases (e.g., political, personal, professional, or religious affiliations). If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the <u>Citation Security Mark</u> should be followed.	<p>SourceDescriptor</p> <p>Classification information is provided in: SourceDescriptor/@ism:*</p>
Source Security Mark	Overall classification marking of the cited resource. As the resource could originate from the US or another country, the <u>Source Security Mark</u> should represent an appropriate US marking or an original non-US marking. The originating country of the classification marking should also be recorded.	<p>ReferencedResourceSecurity</p> <p>The security attributes should represent the contemporary CAPCO marks for the cited resource. The security attributes on the Security element do NOT impact the overall security mark of the citation since they only convey what the classification of the cited resource is not any data in the current document.</p> <p>Additionally if the cited resource has a legacy CAPCO or other marking the text form of that marking could be put in OriginalClassificationMarking Should the OriginalClassification text be in and of itself classified that Classification information may be provided in: OriginalClassificationMarking/@ism:*</p>

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Title (DCMI)	<p>A primary title of the [cited] resource.</p> <p>IC Expansion: There may be multiple titles associated with a given cited resource, especially if the resource is published as part of a larger compilation of materials. Titles associated with a publication, journal, series, or edition may be necessary to uniquely identify the cited resource. Recommended best practice is to provide the cited resource's title, possibly an alternative title if one exists, and to provide additional titles of the larger compilations when necessary. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the <u>Citation Security Mark</u> should be followed.</p>	<p>Title Subtitle CompilationTitle EditionNumber VolumeNumber IssueNumber</p> <p>Classification information is provided in: Title/@ism:* Subtitle/@ism:* CompilationTitle/@ism:*</p>
Type (DCMI)	<p>The nature or genre of the content of the [cited] resource.</p> <p>IC Expansion: The <u>Type</u> includes terms describing general categories, functions, genres, or aggregation levels for content. Examples of <u>Type</u> include publication form (e.g., book, periodical, report, or article), online publication (e.g., Internet site, web page, blog, or wiki), or intelligence discipline (e.g., SIGINT, MASINT, HUMINT). Recommended best practice is to use a controlled vocabulary. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the <u>Citation Security Mark</u> should be followed.</p>	<p>SourceType IntelDiscipline</p> <p>Should the IntelDiscipline be classified it would impact the classification of the overall source citation by changing: IntelDiscipline/@ism:*</p>

2.1.4 Information Security Marking Metadata

The PUBS.XML schemas incorporate the information security marking attributes defined by the *XML Data Encoding Specification for Information Security Marking Metadata* (ISM.XML). These XML attributes are associated with XML elements throughout the model; most often wherever a portion or information resource-level security banner or classification/declassification block is displayed, but also on many other XML elements that may never have their markings displayed.

2.2 Additional guidance

This section provides additional guidance for encoding data in specific situations. In particular, situations for which there is not clearly a single method of encoding the data are documented here. 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 Specification of publishing organization

The element **Publisher** is 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 **Publisher** structure. The publishing organization's approved identifier value is captured in an element called **AgencyAcronym**. The organizations and/or individuals responsible for creating the information are captured in child elements called **AuthorInfo**, **CoauthorInfo**, **ContributingAuthorInfo**, and **POCInfo**. Depending on the enterprise requirement being addressed, a complete understanding of the Publisher requires evaluating the **AgencyAcronym/@acronym** value as well as the values found in the **Affiliation** and **OfficeName** structures of the **AuthorInfo**, **CoauthorInfo**, **ContributingAuthorInfo**, and **POCInfo** elements.

The **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 a publisher uses the required **AgencyAcronym** element. The default CVE for **AgencyAcronym** 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 **OfficeName** element of the **AuthorInfo** or **POCInfo** structures. The **OfficeName** element is officially defined as the name or organizational code of a department, division, branch, or other office within an agency.

For consistency, populate **OfficeName** 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. As one or both of **AuthorInfo** and **POCInfo** are required, the **POCInfo/OfficeName** will have precedent over the **AuthorInfo/OfficeName**, if both are included.

In cases where non-IC information (e.g., OtherDoD, OtherUSG, SLT, Foreign) is shared with the intelligence enterprise, the **AgencyAcronym** 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 **AgencyAcronym** should reflect the appropriate non-IC organization designator and the non-IC organizations office in the appropriate **OfficeName** 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 **OfficeName**, precede the office acronym with the country code tri-graph in order to ensure uniqueness.

Examples

For NCTC:

```
<Publisher><AgencyAcronym acronym="DNI"/>
<POCInfo><OfficeName>NCTC</OfficeName></POCInfo></Publisher>
```

For the XYZ component of NCTC:

```
<Publisher><AgencyAcronym acronym="DNI"/>
<POCInfo><OfficeName>NCTC/XYZ</OfficeName></POCInfo></Publisher>
```

For the XYZ component of CIA:

```
<Publisher><AgencyAcronym acronym="CIA"/>
<POCInfo><OfficeName>XYZ</OfficeName></POCInfo></Publisher>
```

For the United State Postal Service:

```
<Publisher><AgencyAcronym acronym="OtherUSG"/>
<POCInfo><OfficeName>USPS</OfficeName></POCInfo></Publisher>
```

For the JIOC at PACOM:

```
<Publisher><AgencyAcronym acronym="DIA"/>
<POCInfo><OfficeName>PACOM/JIOC</OfficeName></POCInfo></Publisher>
```

For the J4 at PACOM:

```
<Publisher><AgencyAcronym acronym="OtherDoD"/>
<POCInfo><OfficeName>PACOM/J4</OfficeName></POCInfo></Publisher>
```

2.2.2 Specification of Need-To-Know Access Parameters

The element **ntk:Access** is used to identify the parameters needed by a system to allow that system to automatically process an access request involving determination of Need-To-Know. See the *XML Data Encoding Specification for Need-To-Know Metadata* (NTK.XML) for details.

2.2.3 Specification of Production Metrics Reporting Metadata

The element **irm:ProductionMetricsList** is used to identify the metadata needed to support reporting of production metrics as required by DDNI/A. These metadata consists of paired values of **irm:ProductionMetricsCoverage** and **irm:ProductionMetricsSubject**. CVEs for these data elements will be provided as they become available.

2.2.4 Specification of PublicationMetadataList

The element **pubs:PublicationMetadataList** wraps one or more **pubs:PublicationMetadata** elements. For example, the list type may be used to aggregate multiple **pubs:PublicationMetadata** elements of differing classifications, supporting tearline operations. One of the **pubs:PublicationMetadata** could be a TOP SECRET one while another could be SECRET. An example document "RogersRangersMultiMeta.xml" is provided in the Examples directory. The document is UNCLASSIFIED but has example FOUO markings to show having multiple **pubs:PublicationMetadata**. Ensure you comply with [PUBS-ID-00086](#); it requires

classification of the individual **pubs:PublicationMetadata** elements when you have multiple.

2.2.5 Specification of **ism:notice**

The elements **irm:NoticeList**, **pubs:Note**, and **pubs:NoteInline** all have the **ism:notice** attribute on them. There are 3 levels where **irm:NoticeList** is allowed:

- As a child of **pubs:PublicationMetadata**. For example, a FISA notice in this location is used to indicate that the **pubs:PublicationMetadata**, taken alone, would require in and of itself a FISA notice (e.g., when the metadata are displayed as part of a search result set). The FISA notice would not be here if the document required it but the **pubs:PublicationMetadata** did not.
- As a child of **pubs:IntelDoc**. For example, a FISA notice in this location is used to indicate that the warning applies to the document. The **irm:NoticeList** child of **pubs:IntelDoc** must be rendered, but the placement is not defined.
- As a child of **pubs:DocumentBody**. For example, a FISA notice in this location is used to indicate that the **pubs:DocumentBody**, taken alone, would require in and of itself a FISA notice (e.g., when the metadata are displayed as part of a search result set). The FISA notice would not be here if the document required it but the **pubs:DocumentBody** did not.

The elements **pubs:Note**, and **pubs:NoteInline** are allowed in the body of the document and are used to provide a notice near the location of the data requiring a notice.

ISM provides constraint checking in relation to **ism:notice**, requiring that there be a matching between notices used and portions requiring notices. For example, a FISA notice without any FISA portions or vice versa will result in an error or warning, depending on the particular notice.

DoD Distro statements are slightly more complex; a single document may have multiple DoD Distro statements embedded, but may have only one that applies to the whole document. Therefore the appropriate attributes must be applied to the Resource Security Element for the document.

See the example file `SourceCitations.xml` for a sample Notice and use of notice for non-ISM notices.

2.2.6 Specification of approximable dates

Several dates in PUBS.XML including **DateInfoCutoff**, **DateAcquired**, **DateInformation** and **Coverage/Temporal/TimePeriod** use a construct that allows for the date to be expressed in up to three different encodings. The varying encodings allow for degrees of explicitness in both the date value being encoded and the expected processing of the encoded data.

The most unconstrained means of expressing an approximate date is by using a string. The string has no vocabulary associated with it and few if any constraints. The value encoded could be "Sometime last week", "Ramadan 2010", or "Tuesday". As a result of this value being nearly unconstrained, there is little processing expectation beyond search engines indexing the string and allowing it to be searched.

A slightly constrained means of expressing a date is allowing an ISO 8601 date to be specified with an optional attribute **approximation**. The purpose of the attribute is to modify the interpretation of the date with values such as early, late, circa. The ISO 8601 date format as implemented in PUBS.XML allows for differing amounts of specificity such as 2010, 2010-01, 2010-01-01, 2010-01-01T12, 2010-01-01T12:30. No processing behavior is defined for this encoding of dates. The interpretation of when a record marked 2010 should be returned in a query is for each search engine to decide. This includes the interpretation of the values specified in the attribute **approximation**.

It has been reported that some systems have chosen to interpret **DateInfoCutoff=2010** as if it had been 2010-01-01T00:00:00.0Z exactly, while other systems interpret the same value (2010) as matching all dates and times in 2010.

The most constrained encoding of a date is designed to simplify the searching of data that contain date ranges. This involves encoding any date as a range by providing a pair of dates, one representing the start of the range and the other the end of the range. Each of the values is an xsd:dateTime. The elements representing the end points are **EarliestStartDate** and the **LatestEndDate**. It is expected that systems processing dates encoded in this manner should execute queries using a formula equivalent to:

$$\mathbf{EarliestStartDate} \leq \text{QueryValue} < \mathbf{LatestEndDate}$$

Encoding "2010" using an **EarliestStartDate** of 2010-01-01T00:00:00Z and a **LatestEndDate** of 2011-01-01T00:00:00Z would clearly indicate that the date "2010" represents any date-time in 2010.

Using the **EarliestStartDate**, **LatestEndDate** encoding of a date provides the most clarity and gives the producer of the data the greatest control of exactly how that data should be handled by a system.

If the data producer wants to specify the preferred interpretation for either "string" or "approximation" encoded dates, they should also specify the "earliest/latest" values which best encode their intention. It is expected that systems will use the "earliest/latest" for any date-based retrieval while using the other values for display.

2.2.7 Approximable Dates in constraint rules

- When only a string is specified the constraint rules will assume that the date passes any constraints involving it.
- When an ISO 8601 date is specified with varying amounts of precision, the constraint rules will determine if the valid date is a subset of the date specified. For example, 2010 is before 2010-10; since 2010 encompasses all dates in 2010, the valid dates prior to October 2010 are included in that set and the rule would pass. Similarly, 2010 is also after 2010-10 since it encompasses all the dates after October. However, 2010-10 is not before 2010-05 because none of the dates encompassed by 2010-10 occurred prior to 2010-05.
- When earliest/latest values are specified, the rules should still determine if any of the range would satisfy and if so the rule passes.
- When more than one encoding is specified, the most restrictive will be used.
 - Use ISO over String
 - Use earliest /latest over ISO

2.2.8 MIME type

The Multipurpose Internet Mail Extensions (MIME) type for a PUBS.XML document is application/dni-pubs+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 PUBS.XML schema can be found in a separate document entitled *PUBS.XML Schema Guide*. This guide serves as an interactive presentation of the PUBS.XML schema as well as an implementation-specific 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 and HTML files.

Chapter 4 – Data Validation Constraint Rules

Constraint rules explicitly define the validation constraints for PUBS.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. Each constraint rule is defined in an individual file and is encoded in Schematron with an accompanying plain English phrase; however, knowledge of the PUBS.XML schemas and Schematron is required to understand the rules. 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 as defined by the formal Schematron 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 PUBS.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 by IC elements and codified in these constraint rules. Traditional schema languages are generally unable to effectively represent these additional constraints. The constraints can largely be derived from interpreting the intent of the current schemas and/or extracting guidance from the Development Guidance or the policy or guidance documents that govern that type of data.

4.1.1 Schematron

Schematron was selected as the language in which to encode these additional rules. The provided Schematron is used to define the constraint rules; it is NOT a required implementation. Implementers can use any tools at their disposal as long as the data complies with the rules expressed. To facilitate testing and understanding of the rules they are executable in either Oxygen or the XSLT2 implementation of ISO Schematron provided by Rick Jelliffe at <http://Schematron.com>. ISM rules are dependent on XPath 2.0 and XSLT 2.0 features. According to Mr. Jelliffe who is the editor of Schematron for ISO:

“By default, Schematron uses the XPath language as used in XSLT 1.0, and is typically implemented by converting the schema into an XSLT 1.0 script which is run against the document being validated. However, ISO Schematron also allows

XSLT 2.0 to be used, and this is becoming an increasingly popular choice because of the extra expressive convenience of XPath 2.0: a different skeleton is available for this.”

Included in the package are the ISO Schematron implementation XSLT files provided as a convenience along with a compiled version of the rules.

4.1.2 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.

4.1.3 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.

4.1.4 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. PUBS.XML data validation constrain rule IDs are prefixed with “PUBS-ID-”.

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

4.1.5 Errors and Warnings

The severity of a constraint rule violation is categorized as either an “Error” or a “Warning.” An “Error” is more severe and is indicative of a clear violation of a 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 Non-null Constraints

XML syntax allows all elements with content declared to be of data type “string” to have zero or more characters of content — which, allows for empty (or null) content. According to this Specification, all required elements (and certain conditional elements) must have content, other than white space. If an element, defined in this Specification, used in an XML instance is required (or conditional in certain cases), and that element may possibly contain only text content, then the element must have content in order to be Constraint Rules Valid.

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, CVEs and business rules are intended by the author to be used for any particular instance document.

4.4 Inherited Constraints

In an instance of PUBS.XML use of ISM attributes must be fully conformant with the associated version of ISM.XML.

In an instance of PUBS.XML use of ntk:Access must be fully conformant with the associated version of NTK.XML.

4.5 Value Enumeration Constraints

Several elements and attributes of the PUBS.XML model use CVEs to define the data allowed in the element or attribute. In some cases the specific CVE is specified via an attribute, which may include a default CVE. Further, in some of the cases where the CVE can be specified, the attribute may restrict the list of CVEs allowed and some may allow for the author to specify their own CVE. For each of these, the value must be in the specified external CVE or the default CVE.

Some CVEs are not available on all networks. A subset CVE will be provided for use on networks not approved for the entire list. If the processing will occur on a network where the entire CVE is not available, the subset CVE may be substituted in the constraint rules since the excluded values would be excluded from use on the lower network.

As noted in the specific rules, a failure of validation against a CVE will generate an Error.

4.6 Dates and Times

Except for attribute **date** for which the data type is "xsd:date," the data type of each PUBS.XML date/time-related element and attribute is specified as one of four custom simple types defined to allow the full range of allowable patterns specified in the DED for that element or attribute. These four custom simple types are in fact unions of appropriate W3C primitive data types, three of which also include in the union one of two additional custom simple types defined to allow seconds to be optional in time specifications. Schema validation will automatically ensure conformity to the data types. Validations and time comparisons will use the Zulu (Z) time zone when a time zone indicator is absent.

Table 7 summarizes the data types and allowable layout of representations for each of the PUBS.XML date/time-related elements and attributes.

Table 7. Date/Time-Related Data Types and Layout Representations

Element or Attribute	Data Type	Layout of Representation
DateApproved DatePosted DatePublished DateReviewed DateRevised DateValidTil @normalizedDateTime DateTimeReferenced	ISO8601DateTimeType	YYYY(Z ±hh:mm)? YYYY-MM(Z ±hh:mm)? YYYY-MM-DD(Z ±hh:mm)? YYYY-MM-DDThh:mm(Z ±hh:mm)? YYYY-MM-DDThh:mm:ss(.s)?(Z ±hh:mm)?
DateString	common:ShortStringType	A string less than 256 characters such as "Independence day 1980"
DateInformation DateInfoCutoff DateAcquired	ApproximableDateTimeStructureType	A structure consisting of one or more values including a DateString , ApproximableDateTime , ApproximableDateTimeStart , and ApproximableDateTimeEnd ,
ApproximableDateTime ApproximableDateTimeStart ApproximableDateTimeEnd	ApproximableDateTimeType	YYYY-MM-DDThh:mm:ss(.s)?(Z ±hh:mm)? with optional @approximation
EarliestStartDate LatestEndDate	xsd:dateTime	YYYY-MM-DDThh:mm:ss(.s)?(Z ±hh:mm)?
@date	xsd:date	YYYY-MM-DD(Z ±hh:mm)?
@dateTime	dateTimesType	(YYYY(Z ±hh:mm)? YYYY-MM(Z ±hh:mm)? YYYY-MM-DD(Z ±hh:mm)? YYYY-MM-DDThh:mm(Z ±hh:mm) YYYY-MM-DDThh:mm:ss(.s)?(Z ±hh:mm)?)+
@normalizedDate	ISO8601DateType	YYYY(Z ±hh:mm)? YYYY-MM(Z ±hh:mm)? YYYY-MM-DD(Z ±hh:mm)?
@normalizedTime	ISO8601TimeType	hh:mm(Z ±hh:mm) hh:mm:ss(.s)?(Z ±hh:mm)?

Element or Attribute	Data Type	Layout of Representation
@dateTimeRange	dateTimePairsType	((YYYY(Z ±hh:mm)? YYYY-MM(Z ±hh:mm)? YYYY-MM-DD(Z ±hh:mm)? YYYY-MM-DDThh:mm(Z ±hh:mm) YYYY-MM-DDThh:mm:ss(.s)?(Z ±hh:mm)?), (YYYY(Z ±hh:mm)? YYYY-MM(Z ±hh:mm)? YYYY-MM-DD(Z ±hh:mm)? YYYY-MM-DDThh:mm(Z ±hh:mm) YYYY-MM-DDThh:mm:ss(.s)?(Z ±hh:mm)?))+

4.7 Time Zone Indicators

Validations and time comparisons will use the Zulu (Z) time zone, when a time zone is absent. It is recommended that the optional time zone be specified either as Zulu (Z) or as ±hh:mm where applicable.

4.8 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 PUBS.XML

4.9 Obsolete rule numbers

[PUBS-ID-00004] Removed in V4

[PUBS-ID-00005] Replaced by [PUBS-ID-00065]

[PUBS-ID-00006] Replaced by [PUBS-ID-00065]

[PUBS-ID-00010] Replaced by PUBS-ID-00066, PUBS-ID-00067, PUBS-ID-00068, and PUBS-ID-00069.

[PUBS-ID-00012] Replaced by PUBS-ID-00071, PUBS-ID-00072, and PUBS-ID-00073.

[PUBS-ID-00013] Replaced by PUBS-ID-00056, PUBS-ID-00057, PUBS-ID-00058, PUBS-ID-00059, PUBS-ID-00060, and PUBS-ID-00061

[PUBS-ID-00014][Error] Replaced by PUBS-ID-00062, PUBS-ID-00063, and PUBS-ID-00064.

[PUBS-ID-00005][Error] Replaced by [PUBS-ID-00065]

[PUBS-ID-00006][Error] Replaced by [PUBS-ID-00065]

[PUBS-ID-00004] Removed in V4

[PUBS-ID-00077] Removed in V4

[PUBS-ID-00078] Removed in V4

[PUBS-ID-00083] Removed in V4

[PUBS-ID-00084] Removed in V4

[PUBS-ID-00051] Removed in V5

[PUBS-ID-00052] Removed in V5. Replaced by PUBS-ID-00088.

[PUBS-ID-00002] Removed in V6. Replaced by PUBS-ID-00093.

[PUBS-ID-00016] Removed in V6. Replaced by PUBS-ID-00093.

[PUBS-ID-00011] Removed in V6.

[PUBS-ID-00067] Removed in V6.

[PUBS-ID-00069] Removed in V6.

[PUBS-ID-00073] Removed in V6.

[PUBS-ID-00085] Removed in V6.

Chapter 5 – Data Rendering Constraint Rules

The constraint rules in this chapter define constraints on the rendering of PUBS.XML documents. The intent is to inform the development of systems capable of rendering or displaying PUBS.XML data for use by individuals not familiar with the details of the PUBS.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.

5.1 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. PUBS.XML data rendering constraint rule IDs are prefixed with "PUBS-RENDER-"

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

5.2 Errors and Warnings

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

Each system responsible for rendering documents must be evaluated based on its use. Those evaluating the system must make a mission-appropriate decision about the system's suitability for use.

[PUBS-RENDER-00001][Error] When element **RevisionRecall** is present the text of the attribute **revisionType** shall be rendered in uppercase as the first part of the document title immediately following the classification portion mark and will be followed with a ":".

Human Readable: Systems used for rendering data containing the RevisionRecall

element will produce rendered documents that comply with the August 5, 2005 *Negroponete Revision Recall Memo*, and the specific style described.

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

This appendix lists all the acronyms referenced in this DES and lists other acronyms 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 acronyms that are not referenced in that particular DES.

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

DNI – Director National Intelligence

E.O. – Executive Order

GNS – Geographic Names Server

HTML – HyperText Markup Language

IC.ADD – Intelligence Community Abstract Data Definition

IC CIO – Intelligence Community Chief Information Officer

ICD – Intelligence Community Directive

ICS – Intelligence Community Standard

ISBN – International Standard Book Number

ISM – Information Security Marking Metadata

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

NSI – National Security Intelligence

ODNI – Office of the Director of National Intelligence

SSC – Special Security Center

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 4, Edition 1 (Version 4.1). 10 December 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_Manual_FOUO_v4.1_10Dec2010\[1\].pdf](https://www.intelink.gov/sites/ssc/divisions/capco/CAPCO%20Resources/CAPCO_Implementation_Manual_FOUO_v4.1_10Dec2010[1].pdf).

(CAPCO Register)

Authorized Classification and Control Markings Register. Unclassified FOUO version. Volume 4, Edition 1 (Version 4.1). 10 December 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%20Register_FOUO_v4.1_10Dec%202010.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 501)

Director of National Intelligence Chief Information Officer. Intelligence Community Directive Number 501. 21 January 2009. Office of the Director of National Intelligence. http://www.dni.gov/electronic_reading_room/ICD_501.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

(ICS 500-27)

Intelligence Community Standard for Collection and Sharing of Audit Data for IC Information Resources by IC Elements Number 500-27. DRAFT. Office of the Director of National Intelligence.

(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.

(ISO 19757-3:2006)

Information technology -- Document Schema Definition Language (DSDL) -- Part 3: Rule-based validation -- Schematron. 19757-3:2006 International Organization for Standardization (ISO). <http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html>

(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>.

(Schematron Implementation)

<http://www.schematron.com/>

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 8. DES Version Identifier History

Version	Identifier	Date	Purpose
1.0		August 2008	Initial Release
2	ICTechSpec 500.D.3-V2	24 December 2009	Routine revision to technical specification. For details of changes, see appendix F.5
3	ICTechSpec 500.D.3-V3	11 May 2010	Routine revision to technical specification. For details of changes, see appendix F.4
4	ICTechSpec 500.D.3-V4	7 September 2010	Routine revision to technical specification. For details of changes, see appendix F.3
5	ICTechSpec 500.D.3-V5	6 December 2010	Routine revision to technical specification. For details of changes, see appendix F.2
6	ICTechSpec 500.D.3-V6	11 April 2011	Routine revision to technical specification. For details of changes, see appendix F.1

F.1 V6 Change Summary

The following table summarizes the changes made to V5 in developing V6

Table 9. Data Encoding Specification V6 Change Summary

Change	Artifacts changed	Compatibility Notes
Change encoding of constraint rules from text to Schematron.	Documentation Constraint Rules	Other than rules whose changes are noted below this should only result in more clarity of definition for the rules.
Use ISM.XML.V6 and IRM.XML.V4	Schema	Data generation and Ingestion systems need to be updated to properly enforce the new constraint rules.
Remove support for ISO3166 Numeric codes	PUBS-ID-00069 Remove PUBS-ID-00073 Remove CVENumPubsCountryISO3166Numeric Remove CVENumPubsCoverageISO3166Numeric Remove	Data generation and Ingestion systems need to be updated to not use these values anymore and to properly enforce only the remaining constraint rules.
Remove support for ISO3166Digraph codes	PUBS-ID-00067 Remove CVENumPubsCountryISO3166Digraph Remove	Data generation and Ingestion systems need to be updated to not use these values anymore and to properly enforce only the remaining constraint rules.
Remove element DateInformation from PUBS-ID-00054	PUBS-ID-00054 Change	Data generation and Ingestion systems need to be updated to properly enforce the new constraint rule
Replaced PUBS-ID-00002 and PUBS-ID-00016 with PUBS-ID-00093	Documentation PUBS-ID-00002 Remove PUBS-ID-00016 Remove PUBS-ID-00093 Add	Data generation and Ingestion systems need to be updated to use the values. Note: Data valid under previous releases may not be valid under this release.
Element DateString must have content	PUBS-ID-00003 Change	Data generation and Ingestion systems need to be updated to properly enforce the new constraint rule

Change	Artifacts changed	Compatibility Notes
Remove PUBS-ID-00011	PUBS-ID-00011 Remove	Data generation and Ingestion systems need to be updated to properly enforce only the remaining constraint rules.
Update ISO3166Trigraph CVEs	CVEnumPubsCountryISO3166Trigraph Change CVEnumPubsCoverageISO3166Trigraph Change	Data generation and Ingestion systems need to be updated to use the values. Note: Data valid under previous releases may not be valid under this release.
Update ISO3166Digraph CVEs	CVEnumPubsCoverageISO3166Digraph Change	Data generation and Ingestion systems need to be updated to use the values. Note: Data valid under previous releases may not be valid under this release.
Update FIPSDigraph CVEs	CVEnumPubsCountryFIPSDigraph Change CVEnumPubsCoverageFIPSDigraph Change	Data generation and Ingestion systems need to be updated to use the values. Note: Data valid under previous releases may not be valid under this release.
Added support for irm:SubCountryCode to further refine the pubs:CountryCode element	Documentation Schema PUBS-ID-00001 Changed PUBS-ID-00095 Added	Data generation and Ingestion systems need to be updated to use the values and properly enforce the new constraint rules.
Added irm:CountryCode, irm:SubCountryCode, and irm:NonStateActor to check for non-null values	PUBS-ID-00001 Changed	Data generation and Ingestion systems need to be updated to properly enforce the revised constraint rules.
Added support for irm:order attribute to define a sequential ordering of pubs:CountryCode elements	Documentation Schema PUBS-ID-00094 Added	Data generation and Ingestion systems need to be updated to use the values and properly enforce the new constraint rules.
Remove constraints related to compliesWith ICD-710	PUBS-ID-00085 Remove	Data generation and Ingestion systems need to be updated to no longer enforce this constraint.

F.2 V5 Change Summary

The following table summarizes the changes made to V4 in developing V5

Table 10. Data Encoding Specification V5 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
Create Approximable dates.	Documentation Schema	Data generation and Ingestion systems need to be updated to use the new structures. Note: Data valid under previous releases may not be valid under this release.
Create Attachments	Documentation Schema	Data generation and Ingestion systems need to be updated to use the new structures.
Create DateAcquired	Documentation Schema	Data generation and Ingestion systems need to be updated to use the new structures.
Update Mime Types	CVE	Data generation and Ingestion systems need to be updated to use the values. Note: Data valid under previous releases may not be valid under this release.
Update DateInfoCutoff to use Approximable dates	Documentation Schema	Data generation and Ingestion systems need to be updated to use the values. Note: Data valid under previous releases may not be valid under this release.
Create support for CollectionSource	Documentation Schema	Data generation and Ingestion systems need to be updated to use the new structures.
Coverage/Temporal modified to use an approximable date	Schema	Data generation and Ingestion systems need to be updated to use the new structures
Update DateInformation to use Approximable dates	Documentation Schema	Data generation and Ingestion systems need to be updated to use the values. Note: Data valid under previous releases may not be valid under this release.
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.3 V4 Change Summary

The following table summarizes the changes made to V3 in developing V4

Table 11. Data Encoding Specification V4 Change Summary

Change	Artifacts changed	Compatibility Notes
Change to use ISM V4	Documentation PUBS-ID-00085	See ISM change notes for Impacts. Note: Data valid under previous releases will not be valid under this release.
Included irm:ProductionMetricsList	Documentation, Schema	As this is an additional required data element existing data will not be valid and systems will need to be modified to handle the data appropriately.
Eliminate NIPF	Documentation Schema PUBS-ID-00004	Data generation and Ingestion systems need to be updated to properly enforce the new constraint rules
Refactor to single root element IntelDoc	Documentation Schema PUBS-ID-00008	Data generation and Ingestion systems need to be updated to use the new structures.
Use Schema to enforce DES version number	Schema PUBS-ID-00077 PUBS-ID-00078 PUBS-ID-00083 PUBS-ID-00084	Data Ingestion systems need to be updated to use the new Schema instead of constraint rules.
Implemented PublicationMetadataList to allow for multiple metadata descriptions of the document body.	Schema	Data generation and Ingestion systems need to be updated to use the new structures.
Implemented DocumentBody as a container for front and rear matter as well as the ComplexObject Group that contains the main document body.	Schema	Data generation and Ingestion systems need to be updated to use the new structures.
Removed Security Element from DescriptiveMetadata.	Schema	Data generation and Ingestion systems need to be updated to the root node as the Resource security element.
Applied Resource Security to root node IntelDoc.	Schema	Data generation and Ingestion systems need to be updated to the root node as the Resource security element.
Remove references to NIPF	PUBS-ID-00003 PUBS-ID-00004	Data Ingestion systems need to be updated to reflect the rule changes.

Change	Artifacts changed	Compatibility Notes
Require classification of each PublicationMetatdata when there are multiple.	Documentation PUBS-ID-00086	Data generation and Ingestion systems need to be updated to use the new structures.
Add NoticeList	Documentation Schema	Data generation and Ingestion systems need to be updated to use the new structures.
Replace Narcotic with Drug	Schema	Data generation and Ingestion systems need to be updated to use the new structures.
Added irm:RecordKeeper to RecordsManagmentInfo	Schema	As this is an additional required data element existing data will not be valid and systems will need to be modified to handle the data appropriately. Note: this element is only required if its parent, RecordsManagementInfo, is used. Data generation and Ingestion systems need to be updated to use the new structures.

F.4 V3 Change Summary

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

Table 12. Data Encoding Specification V3 Change Summary

Change	Artifacts changed	Compatibility Notes
Change to use ISM V3	Documentation Constraint Rules	See ISM change notes for Impacts. Note: Data could have been created that was valid under previous releases that may not be valid under this release
Update references to E.O. 12958 to refer to E.O. 13526	Documentation	Should not impact data.
Added reference to NTK.XML	Documentation, PUBS-ID-00083	Existing data remain valid. Systems need to be modified to accept and process new data.
Included irm:ProductionMetricsList	Documentation, Schema	As this is an additional required data element existing data will not be valid and systems will need to be modified to handle the data appropriately.
Added constraint to ensure ism:notice is only used where allowed.	PUBS-ID-00082	Existing data remain valid. Systems need to be modified to accept and process new data.

F.5 V2 Change Summary

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

Table 13. Data Encoding Specification V2 Change Summary

Change	Artifacts changed	Compatibility Notes
Various changes to documentation	Documentation Schema	Changes were to correct errors. Any system that relied on the incorrect information may need to be modified.
Removed version number from file names.	Schema	Data generation and ingestion systems need to be updated to use the new file names.
Added ability for instance documents to specify the DES version used for the document.	Schema Constraint Rules	Data generation systems need to be updated to use the new feature. Ingestion systems need to be updated to properly handle the new data.
Changed default namespace.	Schema	Data generation systems need to be updated to use the new namespace. Ingestion systems need to be updated to properly handle the new data.
Added support for Revision/Recall identification.	Schema Constraint Rules	Data generation systems should be updated to use the new structures if they need the feature. Ingestion systems need to use the new specification, including schema and constraints rules. Note: Data could have been created that was valid under previous releases that may not be valid under this release.
Updated reference to DDMS to use version 2.0	Schema	Data generation systems should be updated to use the new structures if they need the feature. Ingestion systems need to use the new specification, including schema. Note: Data could have been created that was valid under previous releases that may not be valid under this release.
Updated DES to support ICD 206 source citations.	Schema Constraint Rules Controlled Value Enumerations	Data generation systems need to be updated to use the new feature. Ingestion systems need to be updated to properly handle the new data. Note: Data could have been created that was valid under previous releases that may not be valid under this release.

Change	Artifacts changed	Compatibility Notes
Relaxed XML Schema for PersonalProfileGroup and replaced with constraint rules.	Schema Constraint Rules	Any system relying solely on the XML schema for validation of PersonProfileGroup may need to be modified to ensure that data are processed appropriately.
Specified and/or update values for numerous elements and attribute to allow more specific validation of data.	Schema Constraint Rules Controlled Value Enumerations	All systems can now use the specified values to ensure data are correct. Note: Data could have been created that was valid under previous releases that may not be valid under this release.

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 CVEs used in this DES are as follows:

Table 14. CVE Definitions

CVE File name	Definition	Attribute and Rules Cross Reference
CVEnumPubsMimeType	All currently valid MIME Types.	@MimeType PUBS-ID-00074 PUBS-ID-00075
CVEnumUNCE20UnitsOfMeasure	All currently valid UNCE-20 units of measure.	@unitOfMeasure PUBS-ID-00062
CVEnumPubsSubjectCodeSystems	All currently valid Subject Code Systems.	@subjectCodeVocabulary SchemaEnumeration
CVEnumPubsIntelSubDisciplineTechniques	All currently valid Intel sub Discipline technique codes.	IntelSubdisciplineTechnique SchemaEnumeration
CVEnumPubsIntelSubDisciplines	All currently valid Intel sub Discipline codes.	IntelSubdiscipline SchemaEnumeration
CVEnumPubsIntelDisciplines	All currently valid Intel Discipline codes.	IntelDiscipline SchemaEnumeration
CVEnumPubsCoverageISO3166Trigraph	WWW and All currently valid ISO-3166 trigraphs	CountryCode PUBS-ID-00068
CVEnumISO639-2Trigraph	All currently valid ISO-639-2 trigraphs.	Language PUBS-ID-00057 PUBS-ID-00060
CVEnumPubsCoverageISO3166Digraph	All currently valid ISO-3166 digraphs.	CountryCode PUBS-ID-00067 PUBS-ID-00059 PUBS-ID-00060 PUBS-ID-00061
CVEnumPubsCoverageFIPSDigraph	All currently valid FIPS digraphs.	CountryCode PUBS-ID-00066
CVEnumPubsCountryISO3166Trigraph	All ISO-3166 trigraphs that have ever been valid	@countryCode PUBS-ID-00072
CVEnumPubsCountryFIPSDigraph	All FIPS 10-4 digraphs that have ever been valid	@countryCode PUBS-ID-00071
CVEnumPubsAgencyAcronym	All currently valid Agency Acronyms for use with publisher.	AgencyAcronym PUBS-ID-00076
CVEnumISO4217Trigraph	All currently valid ISO-4217 Trigraphs.	@unitOfMeasure PUBS-ID-00063
CVEnumISO4217Numeric	All currently valid ISO-4217 Numeric codes.	@unitOfMeasure PUBS-ID-00064
CVEnumISO639Digraph	All currently valid ISO-639-1 Digraphs.	Language PUBS-ID-00056 PUBS-ID-00059 PUBS-ID-00061

CVE File name	Definition	Attribute and Rules Cross Reference
CVEnumISO639-3Trigraph	All currently valid ISO-639-3 Trigraphs.	Language PUBS-ID-00058

Appendix I IRM.HTML to PUBS.XML Mappings

The "Intelligence Community Information Resource Metadata, HTML Encoding" is used by many producing organizations to encapsulate metadata in HTML documents posted to Intelink. In the HTML Implementation, the metadata is stored in **META** elements within the HTML **HEAD** element. This is illustrated by the following example:

```
<HTML>
<HEAD>
<TITLE>(U) Example for Intelink Posting (UNCLASSIFIED)</TITLE>
<META NAME="IL.secur.classif" CONTENT="UNCLASSIFIED">
<META NAME="IL.title" CONTENT="(U) Example for Intelink Posting">
<META NAME="IL.docid" CONTENT="D192-012">
<MEAT NAME="IL.country" CONTENT="AFG">
<META NAME="IL.subcode.ifc" CONTENT="1011">
...
<META NAME="IL.summary" CONTENT="(U) Abstract of this
document...">
</HEAD>
<BODY>
...
</BODY>
</HTML>
```

The HTML Implementation mandates security, product, date, terms-of-use, and electronic (*i.e.*, media format) metadata. The same metadata required by the HTML Implementation is also required by PUBS.XML.

In PUBS.XML, the security metadata is handled by the IC ISM attributes of the root element. For the other metadata categories individual XML elements are provided. For example, "IL.title" corresponds to element **Title** when used in the context that its parent is element **DescriptiveMetadata** and its ancestor is **PublicationMetadata**. Those mapping starting with element **PublicationMetadata** are, in turn, a child of IntelDoc. When metadata from the HTML Implementation maps to an attribute in PUBS.XML, the attribute name, preceded by an "@", is appended to the path. The following path addresses attribute **SCIcontrols** of element **IntelDoc**:
IntelDoc/@SCIcontrols

This appendix contains a mapping of the metadata elements from the HTML Implementation to the corresponding elements and attributes in the PUBS.XML implementation. This is a list of all the actual XML "tags" (XML Element and Attribute *names*) that can appear in an XML document file that maps the Information Resource

Metadata (HTML Encoding) elements to its corresponding Intelligence Community (PUBS.XML) Mapped elements and/or attribute. The PUBS.XML Schema essentially defines a tree structure, containing distinct nodes, each of which can have zero, one, or many associated atomic data values (within an actual schema instance). To explain these mandatory-versus-optional, and singular-versus-repeating constraints, one of the following supplementary annotations has been appended to each tag name.

Table 15. Occurrence codes used in mapping table

Code	Comments
[MS]	Mandatory, Singular (attached to non-repeating, non-optional Elements; also attached to Attributes marked use="required");
[MR]	Mandatory, Repeating (attached to Elements marked maxOccurs="unbounded"; also attached to "list" Attributes, like ownerProducer, which is of type NMTOKENS)
[OS]	Optional, Singular (attached to Elements marked minOccurs="0", but not maxOccurs="unbounded"; also attached to optional Attributes, such as those marked use="optional", but not base="xsd:NMTOKENS")
[OR]	Optional, Repeating (attached to Elements marked minOccurs="0" and maxOccurs="unbounded"; also attached to optional "list" Attributes, such as those marked use="optional" and base="xsd:NMTOKENS"; also attached----rather than using perhaps [XOR]----to any Elements that are listed in an XSD clause of the form: <xs:choice minOccurs="0" maxOccurs="unbounded">, since the listed Elements are neither required nor mutually exclusive in this case)
[MSX]	Mandatory, Singular, but mutually eXclusive with its peers (attached to any Elements that are listed in an XSD clause of the form: <xs:choice>, because these choice structures are subject to the default restrictions minOccurs="1" and maxOccurs="1")
Note that XML Schema notation enables repetition to be accommodated in a wide variety of ways. Our interest here is in the existence of repetition of any kind, due to its profound effect on XML valuation mappings for catalog harvest/ingest operations.	

Table 16. IRM.HTML to PUBS.XML XPath Mappings

IRM.HTML	PUBS.XML XPath Mapping
Security Metadata	
IL.secur.classif	IntelDoc/@classification [MS]
IL.secur.ownerproducer	IntelDoc/@ownerProducer [MR]
IL.secur.ctrl	IntelDoc/@SCIcontrols [OR]
IL.secur.saridentifier	IntelDoc/@SARIdentifier [OR]
IL.secur.fgi	IntelDoc/@FGIsourceOpen [OR] IntelDoc/@FGIsourceProtected [OR]
IL.secur.dissem	IntelDoc/@disseminationControls [OR]
IL.secur.relto	IntelDoc/@releasableTo [OR]
IL.secur.nonlc	IntelDoc/@nonICmarkings [OR]
IL.secur.declassmanualreview	No mapping

IRM.HTML	PUBS.XML XPath Mapping
IL.secur.declasson	IntelDoc/@declassDate [OS] IntelDoc/@declassException[OS] IntelDoc/@declassEvent[OS]
No mapping	IntelDoc/@atomicEnergyMarkings[OR]
No mapping	IntelDoc/@displayOnlyTo[OR]
No mapping	IntelDoc/@derivativelyClassifiedBy[OS]
No mapping	IntelDoc/@derivedFrom[OS]
No mapping	IntelDoc/@compliesWith[OR]
No mapping	IntelDoc/@compilationReason[OS]
No mapping	IntelDoc/@classifiedBy[OR]
No mapping	IntelDoc/@classificationReason[OS]
No mapping	IntelDoc/@nonUSControls[OR]
Product Metadata	
IL.title	PublicationMetadata/DescriptiveMetadata/Title [MS]
IL.docid	PublicationMetadata/AdministrativeMetadata/IdentifierList/DocumentID [OS]
IL.summary	PublicationMetadata/DescriptiveMetadata/Description [MS]
IL.keyword	PublicationMetadata/DescriptiveMetadata/Subject/Keyword [MR]
IL.country	Publisher/AuthorInfo/PostalAddress/CountryCode [OS]
IL.subcode.xxx	PublicationMetadata/DescriptiveMetadata/Subject/SubjectCode [OR]
IL.agency	PublicationMetadata/AdministrativeMetadata/Publisher/AgencyAcronym [MS]
IL.poc	PublicationMetadata/AdministrativeMetadata/Publisher/POCInfo [OR]
IL.itype	PublicationMetadata/DescriptiveMetadata/IntelType [OR]
IL.productline	PublicationMetadata/DescriptiveMetadata/ProductLine [OS]
Date Metadata	
IL.pubdate	PublicationMetadata/AdministrativeMetadata/DateList/DatePublished [MS]
IL.postdate	PublicationMetadata/AdministrativeMetadata/DateList/DatePosted [MSX]
IL.cutdate	PublicationMetadata/AdministrativeMetadata/DateList/DateInfoCutoff [OS]
IL.validtil	PublicationMetadata/AdministrativeMetadata/DateList/DateValidTil [OS]
Terms of Use Metadata	
IL.privacyact	PublicationMetadata/AdministrativeMetadata/Rights/PrivacyActIndicator [MS]
IL.vitalrec	PublicationMetadata/AdministrativeMetadata/RecordsManagementInfo/VitalRecordIndicator [MS]
IL.copyright	PublicationMetadata/AdministrativeMetadata/Rights/CopyRightIndicator [MS]
Electronic Metadata	
IL.applicationtitle	PublicationMetadata/AdministrativeMetadata/RecordsManagementInfo/ApplicationSoftware [MS]
IL.format	PublicationMetadata/DescriptiveMetadata/Description/MediaResource [MS]
IL.url	No Mapping
IL.itype.subdiscipline	PublicationMetadata/DescriptiveMetadata/IntelType/IntelSubdiscipline [OS]
IL.requirement	No Mapping
IL.sensor	No Mapping
IL.reportphase	No Mapping
IL.datetime.range	PublicationMetadata/DescriptiveMetadata/Coverage/Temporal/TimePeriod [MR]
IL.datetime.point	No Mapping

IRM.HTML	PUBS.XML XPath Mapping
IL.analysisistool	No Mapping
IL.designatedactivity	No Mapping