

Please Send Your Comment To : TM-XML-WG@oami.eu.int

TM-XML Version 0.6



DS-XML Version 0.1

XML SCHEMA DESIGN RULES

Version :	0.4
Date :	June 16th, 2005
Document Code :	TM-XML-DR_0.4
Editor :	OHIM XML Working Group

Working Draft

OHIM XML-WG	XML Schema Design Rules		16/06/2010
Version 0.4	Status: Draft		Page 2 of 43

Revisions


Date	Editor	Version	Description
15/04/2004	XML-WG	0.1	Initial draft for TM-XML, a XML standard for Trade Mark.
02/06/2004	XML-WG	0.2	New rules: NM-018, NM-023, NM-024, NM-025, NM-026, DS-006, DS-007, DS-009, DS-010, DS-011, DS-012, DS-015, DS-016, DS-017, DS-018, DS-019.
15/12/2004	XML-WG	0.3	Taking account of comments received in particular from CIPO-Canadian Office and the UBL version 1.0 ratified as OASIS Standard on 08-11-2004. New or updated rules: GE-003, GE-009, GE-010, NM-019, NM-022, NM-027, NM-028, DS-011, DS-020, DS-021. NM-016 blank out because of duplication with NM-022
16/06/2005	XML-WG	0.4	Extension of this document for DS-XML, a XML standard for Design. New rules and new descriptions in the appendix Indicated by the  icon.

Table of Contents

1. INTRODUCTION.....	5
1.1 SCOPE AND PRINCIPAL OBJECTIVES.....	5
1.2 AUDIENCE.....	6
1.3 WORKING GROUP MEMBERS.....	7
1.4 DOCUMENT ORGANIZATION.....	7
1.5 KEY REFERENCES.....	8
1.6 GENERAL CONVENTIONS.....	9
1.7 RULE IDENTIFICATION CODE.....	10
2. GENERAL XML DESIGN RULES.....	11
2.1 CHARACTER SET AND STANDARDS.....	11
2.1.1 <i>Character Set</i>	11
2.1.2 <i>ISO, WIPO, W3C & ebXML Standards Used</i>	11
3. NAMING RULES.....	12
3.1 ISO/IEC 11179 PART 5 NAMING PRINCIPLES FOR DATA ELEMENTS.....	12
3.1.1 <i>Naming Rules Associated to ISO/IEC 11179</i>	12
3.1.2 <i>Naming examples</i>	13
3.2 REPRESENTATION TERMS USED.....	14
3.2.1 <i>Permissible Representation Terms</i>	14
3.2.2 <i>Permissible Representation Terms for Components</i>	15
3.3 OTHER NAMING RULES.....	15
3.4 ACRONYM & ABBREVIATION.....	17
3.5 FILE NAMING RULE.....	17
3.6 AUTHORISED ACRONYMS LIST FOR TAG AND FILE NAMING.....	18
3.7 AUTHORISED ABBREVIATIONS LIST FOR TAG AND FILE NAMING.....	18
4. SCHEMA DESIGN RULES.....	19
4.1 XML DESIGN RULES.....	19
4.2 NAMESPACES.....	20
4.2.1 <i>XML Schema Namespaces</i>	20
4.2.2 <i>XML Instance Namespaces</i>	21
4.3 INCLUDE, IMPORT AND REDEFINE.....	21
5. SCHEMA CONFORMANCE & COMPLIANCE DEFINITION.....	22
5.1 SCHEMA CONFORMANCE LEVEL.....	22
5.2 SCHEMA COMPLIANCE.....	23
6. DOCUMENT ANNOTATION, HEADERS & VERSIONING.....	23
7. EXTENSION OF THE TM-XML BASE VERSION RECOMMENDATIONS.....	24
7.1 EXTENSION & SPECIALIZATION SCENARIOS.....	24
7.1.1 <i>Base and Specialization Scenario</i>	24
7.1.2 <i>Meta-Schema (Union) and Specialization Scenario</i>	25
APPENDIX A : COMPLETE REPRESENTATION TERMS LIST.....	26
APPENDIX B : INTEGRATION WITH EBXML.....	28
B.1. EBXML STACKS.....	28
B.2. EBMS - EBXML MESSAGE SERVICE.....	28
APPENDIX C : RELATED W3C AND OTHERS REFERENCES.....	29
C.1. W3C SOAP 1.1 & 1.2.....	29
C.1.1 <i>SOAP Version 1.2 Description (W3C Recommendation 24 June 2003)</i>	29



OHIM XML-WG	XML Schema Design Rules		16/06/2010
Version 0.4	Status: Draft		Page 4 of 43

C.1.2	W3C SOAP References.	29
C.2.	W3C ENCRYPTION.	30
C.2.1	W3C Encryption Description.	30
C.2.2	W3C Encryption References.	30
C.3.	W3C SIGNATURE.	30
C.2.1	W3C Signature Description.	30
C.2.2	W3C Signature References.	30
C.4.	RASTER GRAPHIC INTERCHANGE STANDARDS.	31
C.1.1	Image Problem Statement.	31
C.1.2	WIPO Image Characteristics.	32
C.1.3	OHIM Image Characteristics.	33
C.1.4	BTO Image Characteristics.	33
C.1.5	DPMA Image Characteristics.	34
C.1.5	UKPO Image Characteristics.	35
C.1.6	JPEG - Joint Photographic Experts Group standard.	36
C.1.8	TIFF - Tag Image File Format standard.	36
C.1.9	PNG - W3C Portable Network Graphics.	37
C.1.10	GIF - Graphic Interchange Format.	37
APPENDIX D : REFERENCES.		38
D.1.	XML DESIGN GUIDE.	38
D.2.	EBXML SPECIFICATION.	38
APPENDIX E : TM-XML SCHEMAS.		40
APPENDIX F : SAMPLE XML DOCUMENTS.		40
APPENDIX G: GLOSSARY & ACRONYMS.		41

1. Introduction.

This guideline is based mainly on commonly used and best practices put into place by some XML Standard initiatives especially UN/CEFACT ebXML, OASIS UBL & xCBL.

The technical challenges are to achieve a consistent and still simple structure with enough flexibility and extension possibilities for any country or Office's specific current and future needs.

However to keep simple and practical, this document is focussed only on the features really necessary and used by the implementation.

Note: This guideline covers only XML Schema which is the format that the Working Group recommends. A XML-DTD can be generated automatically from the XML Schema implementation if needed, but the latter will not be considered as normative.

1.1 Scope and Principal Objectives.

The scope is to create an XML Standard covering the common needs for a Trade Mark Office in the following domains: Trade Mark Applications in particular e-Filing, Publication and Dissemination.

To obtain the objectives defined below, the associated principles are followed:

Objectives	Principles
<p>1</p> <p>Rigour & Independency</p>	<p>Based or aligned to International Standards: W3C Recommendations, ISO 11179 Part 5 for Naming, UN/CEFACT ebXML, OASIS UBL, OMG UML & XMI, etc.</p> <p>Try to be consistent on rules applied. Deviation if exists (for simplification mainly) will be explained and documented.</p> <p>Minimize dependence on any commercial products or proprietary format (Free of intellectual property rights).</p> <p>Preference on use of open source software or freeware when available.</p> <p>Preference on deterministic rules (No ambiguity).</p>
<p>2</p> <p>Simplicity</p>	<p>Avoid additional complexity if there are no real add-values. Any additional analysis or implementation technique introduced should be evaluated beforehand and the add-values should justify the additional complexity.</p> <p>Try to facilitate maintenance and change management. This problem is taken into account from the beginning.</p>

Objectives	Principles
<p style="text-align: center;">3</p> <p>Clarity</p>	<p>Try to find self-described and commonly used vocabulary for naming (Tag, Type, File, etc.) in Oxford English.</p> <p>Specify everything if it brings more clarity. No default rules such as:</p> <ul style="list-style-type: none"> • Type="xs:string" by default if no Type specified <p>Structuring in modular building blocks : AdditionalDatatypes, BusinessTypes.</p> <p>Note: AdditionalDatatypes can be subdivided in AdditionalDatatypes and BusinessEnumerationTypes. BusinessTypes can be subdivided in IPRBusinessTypes and TradeMarkBusinessTypes.</p> <p>Provide Embedded Documentation.</p> <p>Note: The <xs:annotation-documentation> tags will be added at a later stage when the data model description will be stable and close to final. Certainly a link to the Trade Mark Model Documentation, i.e. use of ID features, will be the best solution for element description.</p>
<p style="text-align: center;">4</p> <p>Extensibility & Reusability</p>	<p>As it is not possible to cover all countries/Offices needs (and foresee future needs), it is important to provide an extension feature for the specific and future needs.</p> <p>All Classes are defined as a "Type" and separated into Module-File. The latter can be used as building blocks offering better solutions for Specialization (Restriction, Extension, Re-Use).</p> <p>Use of extension feature below by adding it as last line of all Type:</p> <pre><xs:any namespace="##other" minOccurs="0"/></pre> <p>Offer the possibility of using commercial or standard code list building blocks if desired.</p> <p>Provide examples and recommendations for Office specific extension.</p>

1.2 Audience.

Initial and extension XML Schema designers and developers.

OHIM XML-WG	XML Schema Design Rules		16/06/2010
Version 0.4	Status: Draft		Page 7 of 43






1.3 Working Group Members.





WO	WIPO The World Intellectual Property Organization	Roger Holberton
EM	OHIM Office for Harmonization in the Internal Market (Trade Marks & Designs)	Alexandre Tran
BX	BTO Benelux Trademarks Office	Frank Ten Hulsen
DE	DPMA Deutsches Patent- und Markenamt (German Patent and Trade Mark Office)	Ralf Harmening
GB	UKPO The UK Patent Office	Dave Fraser replaced by John David

1.4 Document Organization.

It is recommended to read the chapters of this guide in order, apart from the appendixes which are supplied for reference or for documenting some useful details.

1.5 Key References.






	References
	XML Schema Part 1 : Structures www.w3.org/TR/xmlschema-1
	XML Schema Part 2 : Datatypes www.w3.org/TR/xmlschema-2
	ISO/IEC 11179-5 Information Technology – Specification and Standardization of Data Elements – Part 5: Naming and Identification Principles for Data Elements.
	ISO 3166-1-Alpha-2 Code Elements (239 Official Country Codes). ISO 639-1:2002 (182 2-Letter Language Codes). ISO 4217-Alpha Currency Codes.
	ISO 8601 - International Standard Date and Time Notation
	ISO/IEC 10646 – Universal Character Set (UCS) – Unicode 3.2 UTF-8.
 	ebXML (Electronic Business Using XML) sponsored by UN/CEFACT and OASIS is a modular suite of specifications for e-Business over the Internet.
<p>World Intellectual Property Organisation</p> 	WIPO ST3: Recommended Standard on 2-Letter Codes for the Representation of States, Other Entities and Intergovernmental Organizations.
	WIPO ST60: Recommendation Concerning Bibliographic Data Relating to Marks.
	International Classification of Goods and Services for the Purposes of the Registration of Marks (Nice Classification).
	International Classification of the Figurative Elements of Marks (Vienna Classification).
	PCT Annex F (Patent Cooperation Treaty).
	WIPO Standard ST.36: Recommendation for the Processing of Patent Documents Using XML.

References	
<p>World Intellectual Property Organisation</p> 	 WIPO ST.80: Recommendation Concerning Bibliographic Data Relating to Industrial Designs.
	 WIPO ST.81: Recommendation Concerning the Content and Layout of Industrial Design Gazettes.
	 International Classification for Industrial Designs (Locarno Classification)








1.6 General Conventions.

We follow the convention defined in the Request for Comment 2119 issued by the Internet Engineering Task Force as summarized below (the icons are an additional feature):

Source: Internet Engineering Task Force (IETF), Request for Comments (RFC) 2219, March 1997. Available at www.ietf.org/rfc/rfc2119.txt?number=2119.

Icon	Term or Phrase Signification
	The term “ MUST ”, “ REQUIRED ” or “ SHALL ” means that the definition is an absolute requirement of the specification.
	The phrase “ MUST NOT ” or “ SHALL NOT ” means that the definition is an absolute prohibition of the specification.
	The term “ SHOULD ” or “ RECOMMENDED ” means that there may exist valid reasons in particular circumstances to ignore a particular item, but the full implications must be understood and carefully weighed before choosing a different course.
	The phrase “ SHOULD NOT ” or “ NOT RECOMMENDED ” means that there may exist valid reasons in particular circumstances when the particular behaviour is acceptable, but the full implications should be understood and the case carefully weighed before implementing any behaviour described with this label.
	The term “ MAY ” or “ OPTIONAL ” means that an item is truly optional.

Other Icons used:

Icon	Description
	To be completed or additional work is required.
	Wait for additional or complementary information.
	Item open to discussion (a doubt exists).
	New rule added to the version (comparing with the previous one).
	Modification done to the version (comparing with the previous one)
	Sunset or obsolete rule.
	Removed

1.7 Rule Identification Code.

The rule identification code is composed of a category prefix of 2 letters followed by a dash as a separator and a sequential number starting from 001 for each category. Any new rules will be assigned the next number sequentially in its category.

Number of rule which is removed in the future will not be re-used and the rule will be kept where it was with the status “Removed” and the removal date.


Note: Nevertheless, before the adoption of the first major version, if necessary, a regrouping, reordering, re-categorizing, renumbering or removal of rules can be done.

Prefix	Description
GE	General schema design rules
NM	Naming rules (File, Tag, Type, Enumeration content, Namespace, etc.)
DS	Design rules (any XSD features)
EB	ebXML integration rules









2. General XML Design Rules.


2.1 Character Set and Standards.

2.1.1 Character Set.

Rule ID	Rule Description	Status
GE-001	ISO/IEC 10646 – Universal Character Set (UCS) – Unicode 3.2 UTF-8 character set must be used.	

2.1.2 ISO, WIPO, W3C & ebXML Standards Used.

Rule ID	Rule Description	Status
GE-002	All XML Schema must be based on the W3C specifications that hold recommendation or proposed recommendation status.	
GE-003	ISO 3166-1-Alpha-2 Code Elements or WIPO ST3 must be used for Country Codes except for the priority and seniority Country –Office which follow the rule GE-006 (Examples: Address Country, Nationality).	
GE-004	ISO 639-1 (2-Letter Language Codes) must be used for Language Codes (Examples: Application, Working, Correspondence Languages).	
GE-005	ISO 8601 - International Standard Date and Time Notation must be used for Date and Time Notation.	
GE-006	WIPO ST3: Recommended Standard on 2-Letter Codes for the Representation of States, Other Entities and Intergovernmental Organizations must be used for Priority and Seniority Country.	
GE-007	The XML Schema should be compliant with UN/CEFACT ebXML (Electronic Business Using XML) and OASIS UBL (Universal Business Language) in particular the ebXML Message Service (ebMS) layer should be used for message packaging, routing and transport facilities.	
GE-008	Regarding the option between XML-DTD and XML Schema, the use of XML Schema is recommended.	
GE-009	ISO 4217-Alpha (3-Letter Currency Codes) must be used for Currency Codes.	



Rule ID	Rule Description	Status
GE-010	<p>For simplicity reason, redefinition of XSD built-in datatypes should be avoided.</p> <p>Example: Use directly the native Boolean type and not an IndicatorType defined as a Boolean or a Boolean type.</p>	



3. Naming Rules.

This chapter addresses file and tag naming conventions which follows with some flexibility the ISO/IEC 11175 Part 5 – Naming principles for data elements.



3.1 ISO/IEC 11179 Part 5 Naming Principles for Data Elements.

3.1.1 Naming Rules Associated to ISO/IEC 11179.

Rule ID	Rule Description	Status
NM-001	Following OASIS UBL and ebXML which derived from ISO 11179-5, the name of the tag must be as much as possible self-described and highly structured.	
NM-002	<p>Tag naming is based on the ISO 11179 Part 5 (Naming and Identification Principles for Data Elements).</p> <p>Element tag name should consist of the Object Class, the name of the Property Term and the name of a Representation Term.</p> <p>An Object Class identifies the primary concept of the element. It refers to an activity or an object within a business context and may consist of one, two or three words.</p> <p>The Property Term identifies the characteristics of the object class. The name of a Property Term shall occur naturally in the tag definition and may consist of one, two or three words. A name of a Property Term shall be unique within the context of an Object Class but may be reused across different Object Classes.</p> <p>If the Representation Term uses the same word as the last one used by the Property Term, the Representation Term shall be omitted.</p>	

Rule ID	Rule Description	Status
NM-027	To express a set of aggregated components or a collection of objects, a Set suffix (as an additional Representation Term) should be used. Examples: TradeMarkSet, SignatorySet.	 

Some flexibility is however accepted for simplification reasons.

Rule ID	Rule Description	Status
NM-003	Object class and Representation Terms should be omitted when the Property Term alone is commonly used and sufficient to express the concept without confusion in its context.	
NM-004	When a same word is repeated in the name, the second or following one(s) should be removed.	

3.1.2 Naming examples.

Examples (Object Class + Property Term + Representation Term):

Object Class	Property Term	Representation Term	Final Name
Applicant	Nationality	Code	ApplicantNationalityCode
GoodsServices	Description	<i>Text</i>	GoodsServicesDescription
<i>Mark</i>	FilingDate	<i>Date</i>	FilingDate

3.2 Representation Terms Used.

A Representation Term categorizes the format of the data element into broad types. A list of UN/CEFACT Representation Terms with some additions is included in Appendix B.

3.2.1 Permissible Representation Terms.

UN/CEFACT Representation Terms	Definition
Code	<p>A character string (letters, figures, or symbols) may for brevity and/or language independence, be used to represent or replace a definitive value or text of a property/attribute. Codes usually are maintained in code lists per attribute type.</p> <p>Examples: CountryCode, CurrencyCode.</p>
Date	A day within a particular calendar year (ISO 8601)
Identifier	<p>A character string used to establish the identity of, and distinguish uniquely, one instance of an object within an identification scheme from all other objects within the same scheme.</p> <p>Note: It shall not be used when an object is identified by its name. In this case “Name” shall be used.</p>
Indicator	A list of two mutually exclusive Boolean values that express the only possible states of a Property. (Values typically indicate a condition such as on/off or true/false.)
Name	A word or phrase that constitutes the distinctive designation of a person, place, thing or concept.
Text	A character string (i.e., a finite set of characters) generally in the form of a word of a language.
Time	The time within a (not specified) day (ISO 8601)









Source: UN/CEFACT, Core Components Technical Specification, Part 1 (Version 2.0), August 11, 2003.








3.2.2 Permissible Representation Terms for Components.

Permissible Representation Terms for Aggregate Core Components or Core Component Types.






Representation Term	Definition
Details	The expression of the aggregation of Core Components to indicate higher levelled information entities.

3.3 Other Naming Rules.

Rule ID	Rule Description	Status
NM-005	All type, element, and attribute names must use Oxford English. Examples: TradeMarkType, ColourClaimedText	
NM-006	Element Name must be in upper camel case (UCC) UCC style capitalizes the first character of each word which compounds the name. Example: AddressCountryCode	
NM-007	ComplexType Name must be in UCC + Suffix Type Example : LanguageCodeType	
NM-008	Attribute must be in lower camel case (LCC). LCC style capitalizes the first character of each word except the first word. Example: <FeeAmount currencyCode ="EUR">	
NM-009	A limit of 35 characters for name is recommended.	
NM-010	Element, attribute and type tag names should be unique.	
NM-011	Tag name should be concise and should not contain consecutive redundant words.	
NM-012	Non-letter characters should not be used.	

Rule ID	Rule Description	Status
NM-013	A tag name and all its components must be in singular form unless the concept itself is plural. Examples: GoodsServices, TotalMarkSeries.	
NM-014	Tag names must only contain nouns, adjectives and eventually verbs. Words like “and”, “of”, “the” must be removed. Example: GoodsServices.	
NM-015	UCC & LCC Name must start with a letter (A-Z) and consist only of those characters included in the NMTOKEN datatype as defined in the W3C Schema definition.	
NM-016	Blank out – duplication of NM-022	
NM-017	Underscores (_), periods (.) and dashes (-) should not be used or should be reserved for some specific usage. Case 1: Country-Office or Domain prefix (XX_UCC with XX in ST3 code). Domain codes will be defined later when necessary. The Country-Office code will certainly be extended with some “Office Union Code” such as mono-class and multi-class system Office code, etc.	
NM-018	The tag name must not be translated, changed or replaced for any purpose.	
NM-019	Regarding the enumeration value or the code list text, it should be short but semantically sufficient and in Oxford English when there is no standard code list. It is a part of the common IP business language. Note: Any translation should be avoided otherwise semantic or cross search or analysis will be more complex. (Translation will be done for the presentation)	 

3.4 Acronym & Abbreviation.

Rule ID	Rule Description	Status
NM-020	Acronyms must not be used unless it is defined in the authorized acronym list for naming. In general it must be commonly used and well-known. Examples: URL, IPR.	
NM-021	The capitalization shall remain when an Acronym is used.	
NM-022	Abbreviations should not be used unless it is defined in the authorized abbreviations list for naming.	
NM-023	When an acronym is defined in the authorized acronym list, it should always be used instead of the complete extended name.	
NM-028	When an abbreviation is defined in the authorized acronym list, it should always be used instead of the complete name.	

3.5 File Naming Rule.

The filename uses a six part naming rule as illustrated below:







N°	Part	Description and syntax
1	[Office]	Use in case of Office specific artifact or union of several offices' code. To be omitted for an office generic version.
2	[Domain]	Indication of the domain. To be omitted for the meta schema
3	Message Flow-Process	Message flow-process.
4	Root Element Message	Root element message.
5	Version	Version and sub-version (separated with the previous part by a dot).
6	Extension	File extension (separated with the previous part by a dot).

Examples:

[Office]	[Domain]	Message Flow-Process	Root Element Message	Version	Extension
EM_	DIS_	Daily	TradeMark	V1-1	.xsd
	EFL_	New	TradeMark	V1	.xsd
EM_		SeniorityClaim	TradeMark	V1	.xsd

Note:

- Field in square brackets [] is optional.
- Associated rules have to be finalized and clarified during technical implementation.

Rule ID	Rule Description	Status
NM-024	Schema and style sheet filename should follow the six part naming rule.	 
NM-025	Message-level schemas should have their versions changed when an included modular schema is updated.	 
NM-026	Filename should follow tag naming rule, however a mapping can be defined locally if the rule cannot be applied due to a technical constraint.	 

3.6 Authorised Acronyms List for Tag and File Naming.









Acronym	Definition
CV2	Card Verification Value (See Appendix G: Glossary & Acronyms for more details)
IPR	Intellectual Property Right
ST3	WIPO Standard 3: Recommended Standard on 2-Letter Codes for the Representation of States, Other Entities and Intergovernmental Organizations.
URI	Uniform Resource Identifier
URL	Uniform Resource Locator

3.7 Authorised Abbreviations List for Tag and File Naming.

Abbreviation	Definition
	None

4. Schema Design Rules.










4.1 XML Design Rules.

Rule ID	Rule Description	Status
DS-001	Restriction on field-length must not be defined.	
DS-002	Element rather than Attribute must be used when the data is subject to extension.	
DS-003	Attributes must be used only for metadata. Example: <GoodsServicesDescription languageCode="es">	
DS-004	Use the <any> element to offer extension leaving the schema opened to additional elements.	
DS-005	XML Schemas should have their versions changed when a referenced external modular schema is updated.	
DS-006	Elements should be declared with occurrence indicators.	
DS-007	The occurrence indicators should not be declared explicitly when the required value is the default value.	
DS-008	The content or value within tags, attributes may be in any language.	

4.2 Namespaces.

4.2.1 XML Schema Namespaces.




Rule ID	Rule Description	Status
DS-009	All XML schemas must declare the W3C schema namespace.	
DS-010	Namespace qualification must be used for W3C schema construct.	
DS-011	The namespace prefixes below must be used : <ul style="list-style-type: none"> ■ xs for W3C schema - XML Schema Definition ■ bt for Business Type ■ at for Additional Type ■ mt for Message Type ■ st for Standard Type 	
DS-012	No default Namespace i.e. both XMLSchema and targetNamespace must be explicitly qualified. This approach, even quite cluttered, is more consistent for all types of schema with no, one or multiple targetNamespaces. Example: <pre><?xml version="1.0"?> <xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema" targetNamespace="http://www.wipo.int" xmlns:lib="http://www.wipo.org" elementFormDefault="qualified"> <xsd:include schemalocation="xxx.xsd"/> </xsd:schema></pre>	
DS-013	Postpone binding schema components to a namespace: Do not give schema a targetNamespace. Let schemas which include another one, which is with no targetNamespace, supply a targetNamespace when it is necessary.	

Rule ID	Rule Description	Status						
DS-014	Hide or expose Namespaces in instance document by using the binary switch attribute: elementFormDefault of the element <xsd:schema> (qualified or unqualified). <table border="1"> <thead> <tr> <th>elementFormDefault</th> <th></th> </tr> </thead> <tbody> <tr> <td>qualified</td> <td>the namespaces will be exposed in instance document</td> </tr> <tr> <td>unqualified</td> <td>Recommendation: namespaces will be hidden (localized) within the schema (not in instance document)</td> </tr> </tbody> </table>	elementFormDefault		qualified	the namespaces will be exposed in instance document	unqualified	Recommendation: namespaces will be hidden (localized) within the schema (not in instance document)	 
elementFormDefault								
qualified	the namespaces will be exposed in instance document							
unqualified	Recommendation: namespaces will be hidden (localized) within the schema (not in instance document)							
DS-015	Schemas must declare a target namespace.							
DS-016	External schema reference should use the “Include” construct. Note: The including and included schemas having the same target namespace.	 						
DS-017	For simplicity, a single-namespace configuration should be preferred. Note: A multiple-namespace may be used for extension purpose (flexibility accepted only during a transition period) - This decision will be done during the technical implementation phase.	 						
DS-018	For clarity, default namespaces must not be used.							
DS-019	All attributes must be namespace qualified.							

4.2.2 XML Instance Namespaces.

Rule ID	Rule Description	Status
DS-		

4.3 Include, Import and Redefine.

Rule ID	Rule Description	Status
DS-020	Include feature is reserved for the inclusion of TM-XML building blocks and Import feature is reserved for Office specific elements, attributes or types importation.	 
DS-021	Redefine technique should be avoided.	

5. Schema Conformance & Compliance Definition.

5.1 Schema Conformance Level.

NEW





Irrelevant		The schema has no feature in common with TM-XML so the question of conformance does not arise.
Consistent		The schema has some feature, in particular the structure, in common with TM-XML and follows the TM-XML Design Rules. (Easy to transform to TM-XML)
Compliant		All features in the TM-XML namespace are valid with TM-XML and all extensions are implemented with other namespace(s).
Conformant		The schema is fully included in the TM-XML so any of its valid instances are also valid with TM-XML.

Note:

Because of first implementations already started, the backward compatibility is now important and will be taken into account, as much as possible, by the Working Group when defining a new version.

5.2 Schema Compliance.

Rule ID	Rule Description	Status
DS-022	<p>To maintain backward schema compliance, a new version may not apply the following changes:</p> <ol style="list-style-type: none"> 1. Rename element or attribute 2. Adding required element or attribute 3. Changing the order of element 4. Removing element or attribute 5. Changing constraint to more restrictive: <ol style="list-style-type: none"> a) Optional to required b) Multiple to single occurrence c) Pattern more restrictive 	 

6. Document Annotation, Headers & Versioning.

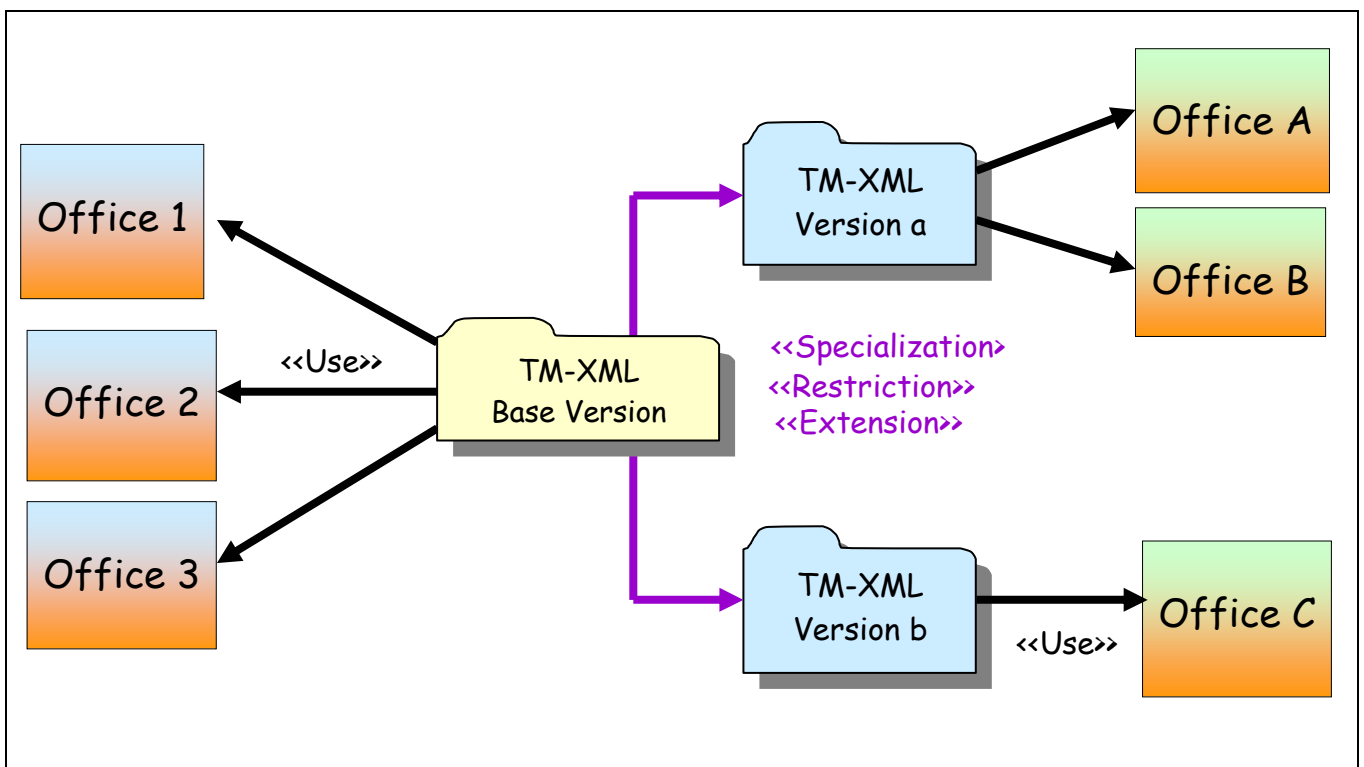
Rule ID	Rule Description	Status
DS-		

7. Extension of the TM-XML Base Version Recommendations.

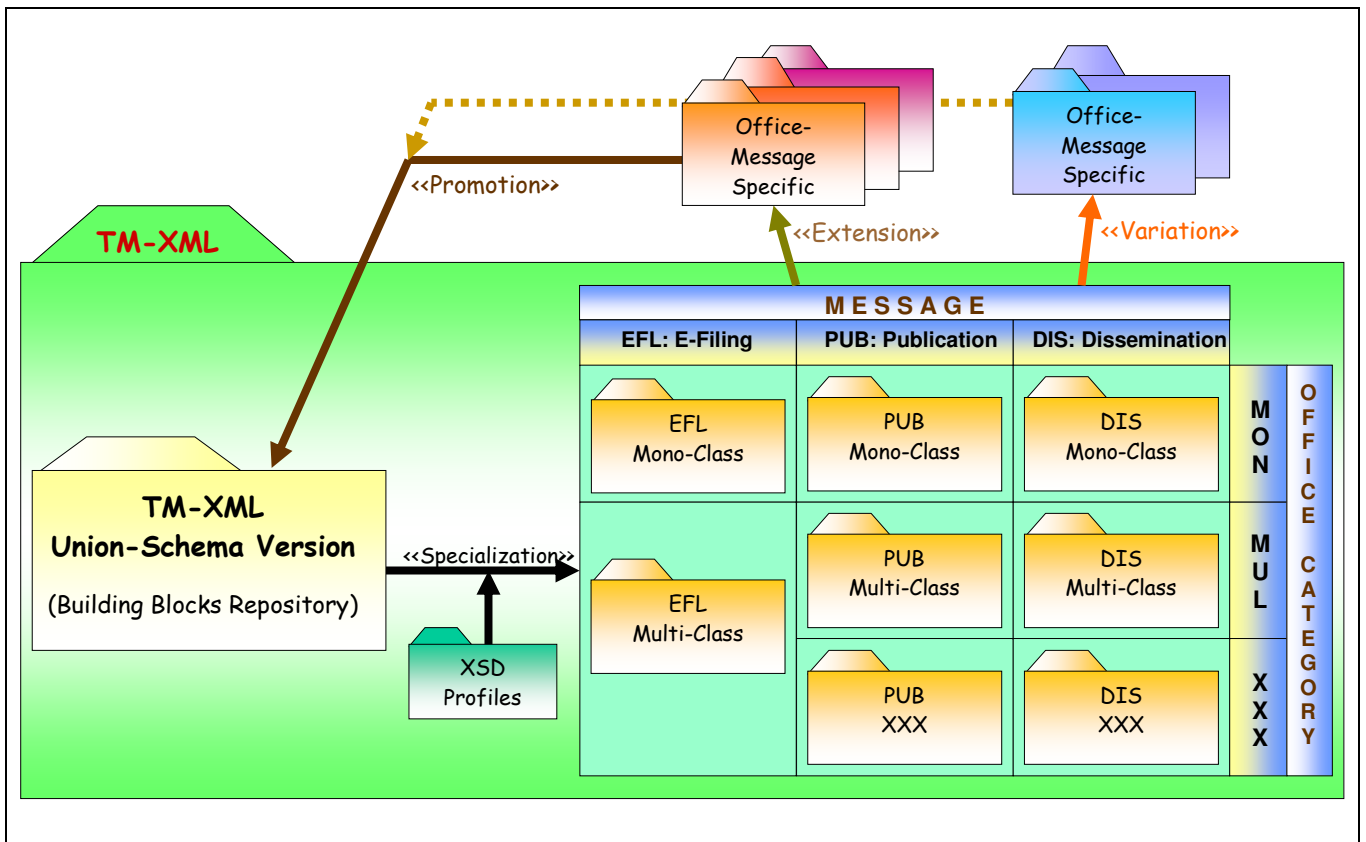
7.1 Extension & Specialization Scenarios.

The objective of the two diagrams below is only for showing the possible approaches. The Office or Messages real needs will be evaluated during the technical implementation phase based on the technical feasibility, the resources available for the maintenance and the feedbacks from Offices.

7.1.1 Base and Specialization Scenario.



7.1.2 Meta-Schema (Union) and Specialization Scenario.



Appendix A : Complete Representation Terms List.

Permissible Representation Terms.

UN/CEFACT Representation Terms	Definition
Amount	A number of monetary units specified in a currency where the unit of currency is explicit or implied.
BinaryObject	A set of finite-length sequences of binary octets. Secondary Representation Terms: Graphic, Picture, Sound, Video
Code	A character string (letters, figures, or symbols) that, for brevity and/or language independence, may be used to represent or replace a definitive value or text of a property/attribute. Codes usually are maintained in code lists per attribute type. Examples: CountryCode, LanguageCode.
Date	A day within a particular calendar year (ISO 8601)
DateTime	A particular point in the progression of time (ISO 8601).
Graphic	A diagram, graph, mathematical curves, or similar representation.
Identifier	A character string used to establish the identity of, and distinguish uniquely, one instance of an object within an identification scheme from all other objects within the same scheme. Note: It shall not be used when an object is identified by its name. In this case "Name" shall be used.
Indicator	A list of two mutually exclusive Boolean values that express the only possible state of a Property. (Values typically indicate a condition such as on/off or true/false.)
Measure	A numeric value determined by measuring an object. Measures are specified with a unit of measure. The applicable unit of measure is taken from UN/ECE Rec. 20.
Name	A word or phrase that constitutes the distinctive designation of a person, place, thing or concept.

UN/CEFACT Representation Terms	Definition
Numeric	Numeric information that is assigned or is determined by calculation, counting, or sequencing. It does not require a unit of quantity or a unit of measure. Secondary Representation Terms: Value, Rate, Percent.
Percent	A rate expressed in hundredths between two values that have the same unit of measure.
Picture	A visual representation of a person, object, or scene.
Quantity	A counted number of non-monetary units possibly including fractions. It is associated with the indication of objects. Quantities need to be specified with a unit of quantity.
Rate	A quantity or amount measured with respect to another measured quantity or amount, or a fixed or appropriate charge, cost or value e.g. kilometre per litre, Euro per class, US Dollars per Euro, etc.
Text	A character string (i.e., a finite set of characters) generally in the form of words of a language.
Time	The time within a (not specified) day (ISO 8601)
Value	Numeric information that is assigned or is determined by calculation, counting or sequencing. It does not require a unit of quantity or a unit of measure.

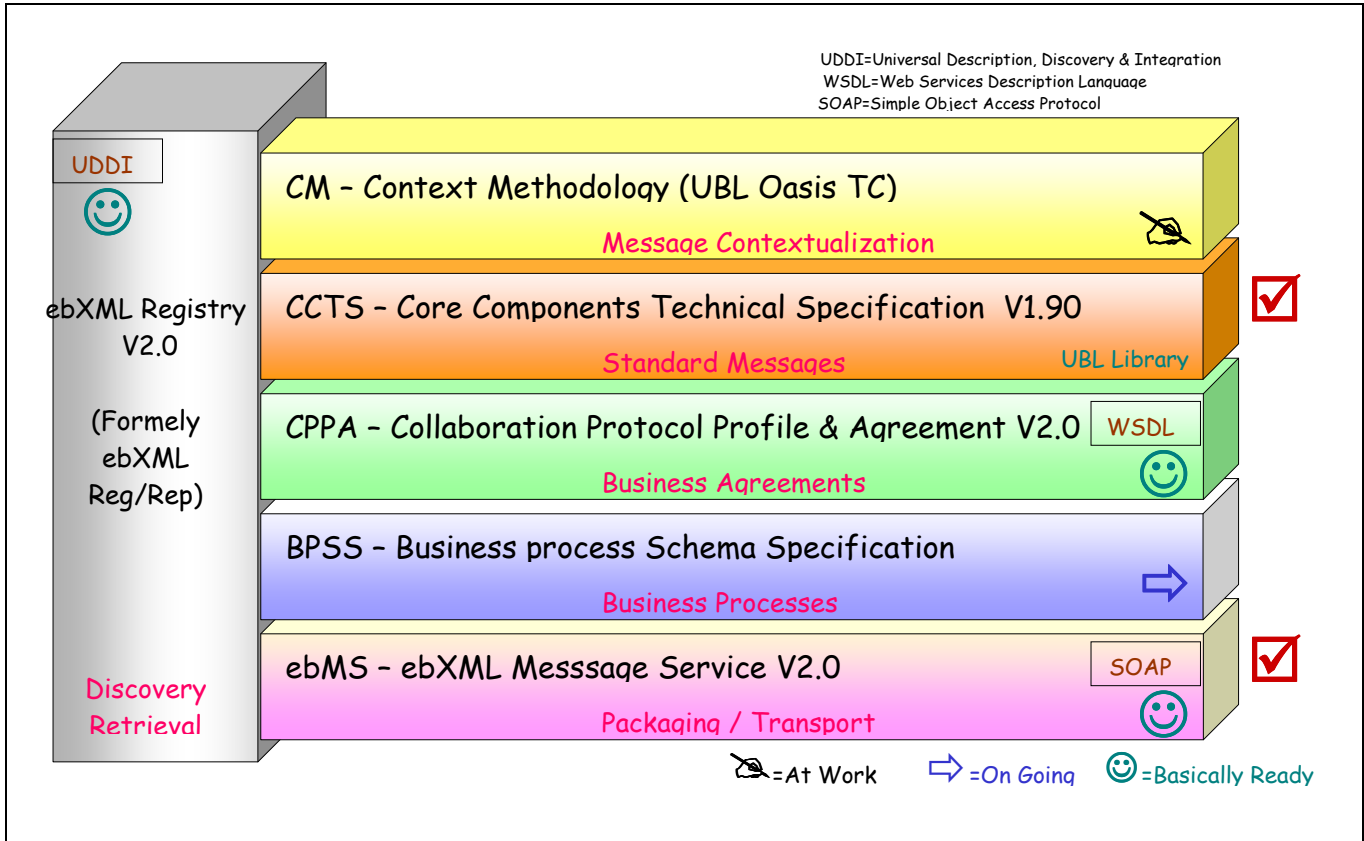
Source: UN/CEFACT, Core Components Technical Specification, Part 1 (Version 2.0), August 11, 2003.

Permissible Representation Terms for Aggregate Core Components or Core Component Types.

Representation Term	Definition
Details	The expression of the aggregation of Core Components to indicate higher levelled information entities.
Type	The expression of the aggregation of Core Components to indicate the aggregation of lower levelled information entities to become Core Component Types. All Core Component Types shall use this Representation Term.
Content	The actual content of an information entity. Content is the first information entity in a Core Component Type.

Appendix B : Integration with ebXML.

B.1. ebXML Stacks.



B.2. ebMS - ebXML Message Service.

Appendix C : Related W3C and Others References.

C.1. W3C SOAP 1.1 & 1.2.

C.1.1 SOAP Version 1.2 Description (W3C Recommendation 24 June 2003).

SOAP, the Simple Object Access Protocol, is a lightweight protocol for exchange of information in a decentralized, distributed environment.

It is an XML based protocol that consists of three parts:

- An envelope that defines a framework for describing what is in a message and how to process it.
- A set of encoding rules for expressing instances of application-defined datatypes.
- A convention for representing remote procedure calls and responses.

SOAP can potentially be used in combination with a variety of other protocols; however, the only bindings defined only describe how to use SOAP in combination with HTTP and HTTP Extension Framework.

C.1.2 W3C SOAP References.

Document	Version - Date
SOAP specification http://www.w3.org/TR/soap12	Version 1.2
SOAP Part0: Primer http://www.w3.org/TR/2003/REC-soap12-part0-20030624/	Version 1.2
SOAP Part1: Messaging Framework http://www.w3.org/TR/2003/REC-soap12-part1-20030624/	Version 1.2
SOAP Part2: Adjuncts http://www.w3.org/TR/2003/REC-soap12-part2-20030624/	Version 1.2
SOAP Specification Assertions and Test Collection http://www.w3.org/TR/2003/REC-soap12-testcollection-20030624/	Version 1.2

C.2. W3C Encryption.

C.2.1 W3C Encryption Description.

C.2.2 W3C Encryption References.

Document	Version - Date
XML Encryption Requirements http://www.w3.org/TR/xml-encryption-req	
XML Encryption Syntax and Processing http://www.w3.org/TR/xmlenc-core/	
Decryption Transform for XML Signature http://www.w3.org/TR/xmlenc-decrypt	
Additional XML Security URIs (Informational): http://www.ietf.org/internet-drafts/draft-eastlake-xmldsig-uri-04.txt	
application/xenc+xml Media Type Registration http://www.ietf.org/internet-drafts/draft-reagle-xenc-mediatype-01.txt	

C.3. W3C Signature.

C.2.1 W3C Signature Description.

XML-Signature Syntax and Processing (W3C Recommendation 12 February 2002).

C.2.2 W3C Signature References.

Document	Version - Date
http://www.w3.org/TR/2002/REC-xmldsig-core-20020212/ http://www.ietf.org/rfc/rfc3275.txt	

OHIM XML-WG	XML Schema Design Rules	16/06/2010
Version 0.4	Status: Draft	Page 31 of 43

C.4. Raster Graphic Interchange Standards.

C.1.1 Image Problem Statement.

In Trade Mark domain, image is an important datatype necessary for information exchange with a certain level of guarantee of its integrity i.e. non degradation, distortion, discoloration or loss of image due to format incompatibility, image compression, colour interpretation or conversion techniques.

The image is created by software or captured by electronic means such as a scanner or a digital camera and represents the logo, picture, shape, etc. for which a protection is claimed.

This artifact can be used for the purposes below:

- Long term storage in the electronic Registry of the original master version
- Display on screen and web browser of any type and resolution
- Resize, conversion or transformation to other version
- Production of bulletin, certificate, letter, etc. by transferring or printing-out on different media such as paper, CD-ROM, DVD, HTML, PDF file, etc.
- Search on the image itself or its metadata
- Image coding such as the Vienna codification or indexation
- Image exchange or transmission with security and reliability
- OCR-ICR or image-shape recognition and comparison

For all these reasons, in particular when exchange of image is required, it is necessary to define standards supported by the industry and IT vendors and sufficiently to guarantee a certain level of integrity, not too complex and at an acceptable cost and time of processing.

Note: Innovation and research are still quite active in this domain. For the moment, there is no real consensus on available formats and versions and there is no one format suitable and accepted for all purposes.

As an example, hereafter are some image characteristics used by the XML-WG member Offices.

C.1.2 WIPO Image Characteristics.

	WIPO	
Format	TIFF	For B&W image
	JPEG	For greyscale and colour image
Resolution	600 DPI	For TIFF image
	266 DPI	For JPEG image
Depth	1 bit	For B&W image
	8 bits	For greyscale image
	24 bits	For colour image
Compression Rate	TIFF	G4
	JPEG	To be defined
Version-Variance	To be defined	
Colour Profile	To be defined	
Remarks	<p>It may be different when images are sent electronically from countries (currently BX, CH, KR, AU, soon US and EM). With these WIPO can have a range from 150 to 350 DPI.</p> <p>WIPO also provide different formats (including GIF, TIFF-G3, BW as JPG/8bits, TIFF-LZW) to different countries on request but try to keep the variations to a minimum.</p>	

C.1.3 OHIM Image Characteristics.

	OHIM
Format	JPEG
Resolution	300 DPI
Depth	24 bits
Compression Rate	(1:8) Not clearly defined yet – Software-Tool dependent. To be defined
Version-Variance	To be defined
Colour Profile	To be defined
Remarks	<p>JPEG is used for the master and as much as possible.</p> <p>GIF version is used for the “light” and “more commonly supported” version on the web. It is produced by a conversion from the master image in JPEG.</p> <p>The image master file is captured with Kofax Ascent Capture software and stored on optical disk jukebox (FileNet Imaging Services).</p>

C.1.4 BTO Image Characteristics.

	BX-BTO (Benelux)
Format	JPEG, GIF
Resolution	300 DPI
Depth	24 bits
Compression Rate	Uncompressed
Version-Variance	To be defined
Colour Profile	To be defined
Remarks	

C.1.5 DPMA Image Characteristics.

DE-DPMA (Germany)							
Format	JPEG						
Resolution	<table border="1"> <tr> <td>300 DPI</td> <td>Scan</td> </tr> <tr> <td>150 DPI</td> <td>Save</td> </tr> <tr> <td>75 DPI</td> <td>Save</td> </tr> </table>	300 DPI	Scan	150 DPI	Save	75 DPI	Save
300 DPI	Scan						
150 DPI	Save						
75 DPI	Save						
Depth	<table border="1"> <tr> <td>8 bits</td> <td>B&W image</td> </tr> <tr> <td>24 bits</td> <td>Colour image</td> </tr> </table>	8 bits	B&W image	24 bits	Colour image		
8 bits	B&W image						
24 bits	Colour image						
Compression Rate	15 %						
Version-Variance	To be defined						
Colour Profile	To be defined						
Remarks	<p>In the field of electronic patent application following image formats are allowed (notice: only black/white or greyscale are allowed for a German patent application) :</p> <table border="1"> <tr> <td>TIFF</td> <td>Resolution: 150 or 300 DPI Depth : 1 or 8 bits Compression: No, LZW or Fax Group 4 Max Size: DIN A4</td> </tr> <tr> <td>JPEG</td> <td>Resolution: 150 DPI Depth: 24 bits Max Size: DIN A4</td> </tr> <tr> <td>PDF</td> <td>Max Size: DIN A4 with fonts embedded</td> </tr> </table> <p>The maximum file size is limited to 2 MB.</p>	TIFF	Resolution: 150 or 300 DPI Depth : 1 or 8 bits Compression: No, LZW or Fax Group 4 Max Size: DIN A4	JPEG	Resolution: 150 DPI Depth: 24 bits Max Size: DIN A4	PDF	Max Size: DIN A4 with fonts embedded
TIFF	Resolution: 150 or 300 DPI Depth : 1 or 8 bits Compression: No, LZW or Fax Group 4 Max Size: DIN A4						
JPEG	Resolution: 150 DPI Depth: 24 bits Max Size: DIN A4						
PDF	Max Size: DIN A4 with fonts embedded						

C.1.5 UKPO Image Characteristics.

UK-PO (United Kingdom)		
Format	TIFF	For B&W, greyscale, RGB colour image
	JPEG	For greyscale and RGB colour image
Resolution	100 DPI	For solid greyscale, spot colour and detailed colour
	150 DPI	For detailed greyscale and colour photo
	600 DPI	For line art B&W graphic
Depth	1 bits	For B&W image
	8 bits	For greyscale and spot colour image
	24 bits	For detailed colour and colour photo image
Compression Rate	TIFF uncompressed	Recommended for solid greyscale, detailed greyscale, detailed colour and colour photo image.
	TIFF-G4	Recommended for line art image
	JPEG	Not defined
Version-Variance	To be defined	
Colour Profile	Image display profile: ICC or sRGB colour space (Red, Green, Blue)	
Remarks	<p>UK-PO follows the image file format accepted by the PaTrAS specification.</p> <p>The PaTrAS specification has also defined rules for the sizing and scaling of the image.</p>	

OHIM XML-WG	XML Schema Design Rules		16/06/2010
Version 0.4	Status: Draft		Page 36 of 43

C.1.6 JPEG - Joint Photographic Experts Group standard.

JPEG is actually the most commonly used format by the trade mark Offices and is a compression of monochrome, greyscale, full-colour or digital still images standard. It uses a 16 million colour palette.

The baseline version uses the lossy, sequential DCT (Discrete Cosine Transform) compression algorithm with Huffman encoding. Typical compression ratios for the lossless algorithm is 8:1 for printed material and up to 35:1 for monitor display.

JPEG is defined by a collection of multi-part International Standards:

ISO/IEC 10918	Information Technology – Digital compression and coding of continuous-tone still images (Part 1 to 4)
ISO/IEC 14495	Information Technology – Lossless and near-lossless compression of continuous-tone still images. (Part 1 & 2)
ISO/IEC 15444	Information Technology – JPEG 2000 Image Coding System (Part 1 to 6)

C.1.8 TIFF - Tag Image File Format standard.

TIFF is a proprietary raster image file interchange format and was developed by Aldus Corporation (owned now by Adobe) and Microsoft and is very widely implemented in DTP (Desktop Publishing) application.

Even if it is recommended as an archival format by the Digital Libraries Federation, there is still a wide variation between requirements and implementations and it is not yet completely harmonized. It can happen that a valid TIFF file cannot be read by a specific version of an imaging software.

Standard TIFF allows the use of different compression techniques such as PackBits, LZW, CCITT Group 3 or 4 Fax and JPEG.

Four photometric classes are supported :

- TIFF-B for monochrome
- TIFF-G for greyscale
- TIFF-P for palette-based coding
- TIFF-R for RGB coding

Latest version: TIFF Revision 6.0

OHIM XML-WG	XML Schema Design Rules	16/06/2010
Version 0.4	Status: Draft	Page 37 of 43

C.1.9 PNG - W3C Portable Network Graphics.

PNG is an extensible file format for the lossless, portable, well-compressed storage of raster images.

PNG provides a patent-free replacement for GIF and can also replace many common uses of TIFF.

Indexed-colour, greyscale, and true color images are supported, plus an optional alpha channel for transparency. Sample depths range from 1 to 16 bits per component (up to 48bit images for RGB, or 64bit for RGBA).

The PNG specification was first issued as a W3C Recommendation on 1st October, 1996 then updated to a second edition incorporating all errata on 10 November 2003. This edition is also an ISO standard, ISO/IEC 15948:2003 (E).

This means it is a mature document that is considered to contribute towards realising the full potential of the Web. The latest version is 1.0.

Viewers for PNG are available on many platforms; there are an increasing number of content creation tools available. Modern browsers implement support for it also.

C.1.10 GIF - Graphic Interchange Format.

GIF is a proprietary raster image file format developed by Compuserve Inc. and widely supported on the Internet. Its latest specification is the version 89a.

The raster data is compressed by the LZW lossless compression algorithm which is patented by IBM and Unisys Corporation.



This format is limited to a 256-colour palette and provides transparency and animation features.

Note:

Following the Unisys attempts to enforce its LZW patent in US, the IETF has developed a replacement encoding, known as the PNG format. JPEG provides a more efficient, standardized, data encoding algorithm.

Appendix D : References.

D.1. XML Design Guide.

Document	Version - Date
Draft Federal XML Developer's Guide U.S. Federal CIO Council xml.gov/documents/in_progress/developersguide.pdf	April 2002
DON XML Developer's Guide US Department of the Navy	Version 1.1 1 May 2002
e-Government Schema Guidelines for XML Office of the e-Envoy, UK Online www.e-envoy.gov.uk/assetRoot/04/00/09/75/04000975.pdf	Version 3.1 January 2004
XML Design Rules and Conventions for the Environmental Information Exchange Network - US Environmental Protection Agency	Version 1 September 2003
 X12 Reference Model for XML Design www.x12.org/x12org	2002-10
 HK XML Schema Design and Management Guide (4 parts) www.itsd.gov.hk/itsd/english/infra/eif.htm .	2003

D.2. ebXML Specification.

Document	Version - Date
ebXML Glossary www.ebxml.org/specs/ebGLOSS.pdf	Version 1.0 May 2001
ebXML Technical Architecture www.ebxml.org/specs/ebTA.pdf	Version 1.0.4 16 February 2001
ebXML Business Process Specification Schema www.ebxml.org/specs/ebBPSS.pdf	Version 1.01 11 May 2001

OHIM XML-WG	XML Schema Design Rules	16/06/2010
Version 0.4	Status: Draft	Page 39 of 43

Document	Version - Date
Message Service Specification OASIS ebXML Messaging Services Technical Committee www.oasis-open.org/committees/ebcml-msg/documents/ebMS_v2-0.pdf	Version 2.0 21 February 2002
OASIS Universal Business Language (UBL) (Ratified as an OASIS Standard on 8 th November 2004.) docs.oasis-open.org/ubl/cd-UBL-1.0/ docs.oasis-open.org/ubl/cd-UBL-1.0.zip	Version 1.0 15 September 2004

Appendix E : TM-XML Schemas.

All of TM-XML schemas and building blocks are stored in the following directories:

Directory	code	Content
AdditionalType	at	All additional types
BusinessType	bt	All business types
CountrySpecificType	ct	Country specific types
MessageType	mt	Message and transaction types
StandardType	st	Standard enumeration types

For convenience in printing the schema or searching of a word, an all-in-one schema per message is supplied.

Schema dependencies among schema modules, components and building blocks will be studied and dependency diagram and recommended structure will be supplied in the version 1.0.

Appendix F : Sample XML Documents.

The XML document will be supplied in separate files and an example will be integrated in this appendix in later version.

Appendix G: Glossary & Acronyms.

Term	Definition
BPSS	Business Process Schema Specification (an ebXML stack).
CCITT	Consultative Committee on International Telegraphy and Telephony, an organization that sets international communications standards renamed in 1993 as ITU (the parent organization).
CCTS	Core Components Technical Specification (an ebXML stack).
CD-ROM	Compact Disk Read-Only Memory.
Compositor	A W3C XML Schema construct that groups element declarations : Sequence, Choice, All.
CPPA	Collaboration Protocol Profile and Agreement (an ebXML stack).
CV2	<p>CV2 is a Card Security Code (CSC), basically the 3 or 4 digit number located on the back of a credit/debit card signature strip.</p> <p>It is a collective term from:</p> <ul style="list-style-type: none"> • Visa's CVV2 – Card Verification Value • MasterCard 's CVC2 – Card Validation Code (CVC2) • American Express's CID – Card Identification Number
DPI	Dots Per Inch. A measure of image resolution for printers, scanners and displays.
DTD	Document Type Definition.
DVD	Digital Versatile Disk (or formerly for film storage: Digital Video Disk).
ebMS	ebXML Message Service (an ebXML stack).
ebXML	Electronic business XML.
ICC	<p>The International Color Consortium (see www.color.org) was established in 1993 by 8 industry vendors for the purpose of creating, promoting and encouraging the standardization and evolution of an open, vendor-neutral, cross-platform colour management system architecture and components.</p> <p>The outcome of this cooperation was the development of the ICC profile specification. The current version is the ICC 1:2003-09 File Format for Colour Profile Version 4.1.0.</p>

Term	Definition
ICR	Intelligent Character Recognition (used mainly for no printed characters).
IPR	Intellectual Property Right.
ITU	International Telecommunication Union.
LCC	Lower Camel Case. A naming convention that capitalizes the first character of each word except the first word and compounds the name. Example: languageCode
LZW	A compression/decompression algorithm created by Lempel Abraham, Ziv Jakob and Welch Terry (Unisys Corporation).
OASIS	OASIS (Organization for the Advancement of Structured Information Standards) is non-for-profit, international consortium that drives the development, convergence and adoption of e-business standards.
OCR	Optical Character Recognition.
PDF	Portable Document Format.
RFC	Request for Comments.
SOAP	Simple Object Access Protocol (See Appendix C)
UBL	Universal Business Language (OASIS).
UCC	Upper Camel Case. A naming convention that capitalizes the first character of each word and compounds the name. Example: ApplicationLanguageCode
UCS	Universal Character Set.
UDDI	Universal Description, Discovery and Integration.
UN/CEFACT	United Nations Centre for Trade Facilitation and Electronic Business.
URI	Uniform Resource Identifier.
URL	Uniform Resource Locator.
UTF	UCS transformation format.
W3C	World Wide Web Consortium.

Term	Definition
WSDL	Web Services Description Language.
xCBL	The XML Common Business Library is a set of XML building blocks and a document framework that allows the creation of robust and reusable XML documents to facilitate global trading.
XML	Extensible Markup Language.
XSD	Schema Definition Language (Synonym: XML Schema).