



STANDARD ST.66
DECEMBER 2007 CHANGES

	<u>Pages</u>
REFERENCES	2
Editorial changes.....	2
REQUIREMENTS OF THE STANDARD.....	3
Paragraph 17, revised November 2007	3
Paragraph 22, revised November 2007	3
Paragraph 23, revised November 2007	3
Characters - Paragraph 28, revised November 2007	3
Characters - Paragraph 29, revised November 2007	3
General XML Constructs - Paragraph 40, revised November 2007	4
General XML Constructs - Paragraph 41, revised November 2007	4
Naming conventions - Paragraph 47, revised November 2007.....	5
Naming conventions - Paragraph 64, revised November 2007.....	7
Miscellaneous XSD Rules - Paragraph 69, added November 2007.....	7
Paragraph 70, added November 2007.....	7
External Entities - Paragraph 76, revised November 2007.....	8
External Entities - Paragraph 77, added November 2007	8



STANDARD ST.66

Version 1.1

RECOMMENDATION FOR THE PROCESSING OF TRADEMARK INFORMATION USING XML
(EXTENSIBLE MARKUP LANGUAGE)

*Revision adopted by the ST.66 Task Force of the Standards and Documentation
Working Group (SDWG) on December 3, 2007*

TABLE OF CONTENTS

WIPO STANDARD ST.66	1
INTRODUCTION.....	1
DEFINITIONS AND TERMINOLOGY	1
SCOPE OF THE STANDARD.....	2
REFERENCES.....	2
REQUIREMENTS OF THE STANDARD	3
Characters	3
General XML Constructs	4
Naming and Modeling Constraints.....	4
Reusability Scheme	4
Namespace Scheme.....	4
Naming Conventions	5
XML Tag Naming Rules.....	5
Acronyms and Abbreviations	6
XML Schema File naming Rules	6
Miscellaneous XSD Rules	7
Naming Office-Specific Types and Elements.....	7
External Entities.....	8

APPENDICES

The Appendices are available at: http://www.wipo.int/standards/en/xml_material/st66/

- Appendix A – ST.66 XML Dictionary
- Appendix B – ST.66 XML Schema
- Appendix C – ST.66 Associated Class Diagram
- Appendix D – List of Acronyms and Abbreviations
- Appendix E – ST.36 Compatibility Form



WIPO STANDARD ST.66

Version 1.1

RECOMMENDATION FOR THE PROCESSING OF TRADEMARK INFORMATION USING XML (EXTENSIBLE MARKUP LANGUAGE)

*Revision adopted by the ST.66 Task Force of the Standards and Documentation
Working Group (SDWG) on December 3, 2007*

INTRODUCTION

1. This Standard recommends the XML (eXtensible Markup Language) resources used for filing, processing, publication, and exchange of all types of trademark information. It is based in large part on TM-XML, Trade Mark Model at <http://www.tm-xml.org/>, WIPO Madrid Electronic CommunicAtions (hereafter referred to as MECA) DTD and WIPO Standard ST.36.
2. The term "XML resources" is intended to refer to any of the components used to create and operate an XML implementation. For further information about the W3C (World Wide Web Consortium), see <http://www.w3c.org/>.
3. The term "XML Schema" is a language for describing the structure and constraining the contents of XML documents.
4. There are many schema languages based on XML. This Standard recommends only the W3C XML Schema language. The term "XML Schema Definition (XSD)" is an instance of an XML schema written in the W3C's XML Schema language. An XSD defines a class of XML document instances in terms of constraints upon what elements and attributes may appear, their relationship to each other, what types of data may be in them, etc.
5. XML cannot be used *per se* as the basis for trademark document processing. Therefore, this Standard defines elements and their generic identifiers, or "tags", and attributes for marking up trademark documents. That is, this Standard provides for some level of the semantics (meaning), the use, and the names of the types, elements and attributes that make up the various document types it discusses.
6. The purpose of the Standard is to provide logical, system-independent structures for trademark document processing, whether for text or image data. This Standard refers to ISO Standards for country code (ISO3166), language code (ISO639), currency code (ISO4217) and ST.3 code as external schema.

DEFINITIONS AND TERMINOLOGY

7. The keywords MUST, MUST NOT, SHALL, SHOULD, SHOULD NOT, MAY, and OPTIONAL, when they appear in this document, are to be interpreted as described in Internet Engineering Task Force (IETF) Request For Comments (RFC) 2119. Non-capitalized forms of these words are used in the regular English sense.
 - (a) Example – A representation of a definition or a rule. Examples are informative.
 - (b) [Note] – Explanatory information. Notes are informative.
8. For the purpose of this recommendation the expression:
 - (a) "mark" means a trademark, a service mark or another type of distinguishing sign according to the definition of the mark in the legislation concerned, including but not limited to collective marks, certification marks or guarantee marks;
 - (b) "certificate" means the official document which is delivered to a mark owner certifying that his/her mark has been registered by the Office of the country or organization in question, or that such registration has been renewed or amended (this definition also covers "certificates" or "registry extracts" delivered by the Office, e.g., for the purposes of proceedings in court);
 - (c) "gazette" means an official publication containing announcements relating to marks and made in accordance with requirements under national industrial property laws or international industrial property conventions or treaties;



HANDBOOK ON INDUSTRIAL PROPERTY INFORMATION AND DOCUMENTATION

Ref.: Standards - ST.66

page: 3.66.2

(d) "INID" is an acronym for "Internationally agreed Numbers for the Identification of (bibliographic) Data."

9. Markup is defined as text that is added to the content of a document and that describes the structure and other attributes of the document in a non-system-specific manner, independently of any processing that MAY be performed on it.

10. For other definitions see the XML specification at <http://www.w3c.org/TR/2004/REC-xml11-20040204/>.

SCOPE OF THE STANDARD

11. This Standard is aimed at providing guidance to national, regional and international authorities who, on the basis of national industrial property laws or international industrial property conventions, publish announcements, either on trademark applications or on registrations of trademarks.

12. This Standard is intended to provide XML resources for exchanging and processing trademark documents and trademark transaction records. See Appendix 2 for model schemas related to various types of trademark documents and transaction records.

13. This guideline covers only W3C XML Schema. Although an XMLDTD can be generated automatically from an XML Schema that does conform to this Standard, by definition, no DTD can conform to this Standard. See <http://www.w3.org/XML/Schema#dev> for the XML Schema specification.

REFERENCES

14. The following standards and documents are of relevance to this Standard:

- (a) WIPO Standard ST.3: *Recommended Standard on two-Letter Codes for the Representation of States, Other Entities and Intergovernmental Organizations*;
- (b) WIPO Standard ST.36: *Recommendation for the Processing of Patent Documents Information Using XML*;
- (c) WIPO Standard ST.60: *Recommendation Concerning Bibliographic Data Relating to Marks*;
- (d) WIPO Standard ST.62: *Recommendation Concerning Standard Abbreviation for "Vienna Classification"*;
- (e) WIPO Standard ST.63: *Recommendation Concerning the Content and Layout of Trademark Gazettes*;
- (f) WIPO Standard ST.64: *Recommended Search Files for Trademark Search*;
- (g) International Classification of Goods and Services for the Purposes of the Registration of Marks (Nice Classification);
- (h) International Classification of the Figurative Elements of Marks (Vienna Classification);
- (i) TM-XML (<http://www.tm-xml.org/>) version 1.0: *Specification, Dictionary and Design Rules*;
- (j) ISO/IEC 11179-5 *Information technology – Metadata registries (MDR) – Part 5: Naming and identification principles*;
- (k) ISO 3166-1 – *Codes for the representation of names of countries and their subdivisions – Country Codes*;
- (l) ISO 639-1 – *Codes for the representation of names of languages – Part 1: Alpha2-code*;
- (m) ISO 4217 – *Codes for the representation of currencies and funds*;
- (n) ISO 8601 – *Data elements and interchange formats – Information interchange – Representation of dates and times*;
- (o) ISO/IEC 10646 – *Information technology – Universal Multiple-Octet Coded Character Set (UCS)*;



HANDBOOK ON INDUSTRIAL PROPERTY INFORMATION AND DOCUMENTATION

Ref.: Standards - ST.66

page: 3.66.3

(p) ebXML (Electronic Business using XML) sponsored by UN/CEFACT (United Nations Centre for Trade Facilitation and Electronic Business) and OASIS (Organization for the Advancement of Structured Information Standards) is a modular suite of specifications for e-Business over the Internet.⁽¹⁾

(q) UN/CEFACT- XML Naming and Design Rules, Version 2.0;

(r) OASIS UBL Naming and Design Rules;

(s) Internet Engineering Task Force (IETF) Request For Comments (RFC) 2119.

REQUIREMENTS OF THE STANDARD

15. ST.66 XML Dictionary in Appendix A is the foundation of this Standard.

16. The Dictionary MUST be used as defined in this Standard, that is, the types, elements, attributes and enumerations MUST be as indicated in the list of the Dictionary. However, some enumerations are defined as open and may be restricted or extended in the specific offices' implementation.

17. Implementation that conforms to this Standard MUST be carried out built according to the guidelines in this Standard, or MUST be an extension of a conforming XSD according to the guidelines in this Standard.

18. XML instances that conform to this Standard MUST be well-formed XML and validated by the XSD in Appendix B.

19. It is understood that this Standard cannot possibly include all elements required by all trademark offices; in such implementation schema, office-specific elements are allowed as described below in the "Naming Office-Specific Types and Elements" section.

20. The W3C XML Schema definition language has become the generally accepted schema language that is experiencing the most widespread adoption. Although other schema languages exist that offer their own advantages and disadvantages, all XML Schema design rules MUST be based on the W3C XML Schema Recommendations: XML Schema Part 1: Structures and XML Schema, Part 2: Datatypes. All schemas and messages MUST be based on the W3C suite of technical specifications holding recommendation status.

21. Redefinition of XSD built-in data types SHOULD be avoided.

22. WIPO Standard ST.3 MUST be used for priority, ~~and~~ seniority country and designated country.

23. ISO 3166 MUST be used for address country codes, exhibition country codes and nationalities.

24. ISO/IEC 10646 – UCS – Unicode UTF-8 MUST be used for character set.

25. ISO 639-1 (2-Letter Language Codes) MUST be used for Language Codes.

26. ISO 8601 - International Standard Date and Time Notation MUST be used for Date and Time Notation. W3C Schema data types include date and time and SHOULD be used in preference to ISO 8601, where there is any conflict.

27. ISO 4217-Alpha (3-Letter Currency Codes) MUST be used for Currency Codes.

Characters

28. This Standard recommends Unicode exclusively, ~~even though XML permits other character encodings.~~ It may be useful to add character entities for characters not yet in Unicode, ~~such as those listed in wipo.ent (located at <http://www.wipo.int/pct-safe/opct/schemaDocs/1.4/schemaDocs.htm>).~~ This entity file provides general entity names that ~~can be used in instances in place of the code points from the encodings that they are mapped to in wipo.ent.~~ Use of these entities requires the creation of glyphs for presentation, which do not yet exist. See <http://www.w3.org/XML/Core/2002/10/charents-20021023> for further information about character entities.

29. Document instances MUST include ~~an XML declaration~~ ~~the following processing instruction~~ as the first line in the file.

⁽¹⁾ Editorial notes: ebXML was published in 1999 as an initiative of the [United Nations Centre for Trade facilitation and Electronic Business](#) (UN/CEFACT) and the [Organization for the Advancement of Structured Information Standards](#) (OASIS).



HANDBOOK ON INDUSTRIAL PROPERTY INFORMATION AND DOCUMENTATION

Ref.: Standards - ST.66

page: 3.66.4

```
<?xml version='1.0' encoding='utf-8' ?>
```

Note that only UTF-8 is recommended supported in this Standard.

However, in the case of ideographic scripts, Unicode in UTF-8 MAY produce exceptionally large files since the encoding MAY use up to four bytes per character. In such cases, national Offices MAY select an encoding that brings files to manageable sizes. Offices that elect to do so, SHOULD be prepared to consult with their exchange partners and to give adequate public notice.

30. The characters that are permitted to appear in an XML document are specified in the *XML 1.0 W3C Recommendation*, and are endorsed by this Standard with the following exception. The characters used in type, element, or attribute names described in this Standard are restricted to the following set: {a-z, A-Z, 0-9, period (.), dash (-) and underscore (_)}.

General XML Constructs

Naming and Modeling Constraints

Naming Constraints

31. Each dictionary entry name MUST define one and only one fully qualified path for an element or attribute.

Modeling Constraints

32. Libraries and Schemas MUST only use approved datatypes.

33. Mixed content MUST NOT be used in data centric schema except where contained in an `xsd:documentation` element.

Reusability Scheme

34. All type declaration MUST be global.

Namespace Scheme

Declaring Namespaces

35. Every schema module, except internal schema modules, MUST have a namespace declared using the `xsd:targetNamespace` attribute.

36. All XML schemas MUST declare the W3C schema namespace. Schemas MUST declare a target namespace.

37. Namespace qualification MUST be used for W3C schema construct.

38. Every defined or used schema set version MUST have its own unique namespace.

39. Published namespaces MUST never be changed.

40. There SHALL be no default namespaces. That is, for example, both the `XMLSchema` and `targetNamespace` MUST be explicitly qualified. This approach, even though quite cluttered, is more consistent for all types of schema with no, one, or multiple targetNamespaces. An For example is as below: -

```
<?xml version="1.0"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema"
            targetNamespace="http://www.wipo.int/standards/XMLSchema"
            xmlns:lib="http://www.wipo.int/standards/XMLSchema"
            elementFormDefault="qualified">
<xsd:include schemalocation="xxx.xsd"/>
</xsd:schema>
```

41. - To hide or expose Namespaces in instance documents, use the binary switch attribute SHOULD be used: `elementFormDefault` of the element `<xsd:schema>` (qualified or unqualified).

42. External schema references SHOULD use the "Include" construct. The including and included schemas MUST have the same target namespace.

43. For simplicity, a single namespace configuration SHOULD be preferred. Multiple namespaces MAY be used for extension purposes (nationalization).



Naming Conventions

XML Tag Naming Rules

44. The XML tag naming conventions are based on the concepts as defined in ISO 11179 Part 5. Element, attribute and type names SHOULD consist of the Object Class, the name of the Property Term and the name of a Representation Term.

(a) 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.

(b) 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.

(c) If the Representation Term uses the same word as the last one used by the Property Term, the Representation Term SHALL be omitted.

(d) 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.

(e) For example (Object Class + Property Term + Representation Term):

- ApplicantNationalityCode: Applicant(Object Class) + Nationality(Property Term) + Code(Representation Term)
- GoodsServicesDescription: GoodsServices(Object Class) + Description(Property Term) + Text(Representation Term, omitted)
- FilingDate: Mark(Object Class, omitted) + FilingDate(Property Term)+ Date(Representation Term, omitted)

45. Element, attribute and type names MUST be unique. The names SHOULD be concise and SHOULD NOT contain consecutive redundant words, and MUST be as much as possible self-described and highly structured.

46. Element, attribute and type names and all their components MUST be in singular form unless the concept itself is plural. For example, GoodsServices, TotalMarkSeries.

47. Element, attribute and type names MUST only contain nouns, adjectives and eventually verbs. Words like "and", "of", "the" MUST be removed, **except when this makes the name misleading**. For example, GoodsServices, DocumentSizeInByte.

48. Element, attribute and type names MUST NOT be translated, changed or replaced for any purpose.

49. Element, attribute and type names MUST be composed of words in the English language, using the primary English spellings provided in the Oxford English Dictionary, including office-specific tags (except see paragraph 57 below for acronyms).

50. Element names MUST be in upper camel case (UCC). UCC style capitalizes the first character of each word which compounds the name. For example, AddressCountryCode.

51. Type names MUST be in UCC + Suffix Type. For example, LanguageCodeType.

52. Attribute names MUST be in lower camel case (LCC). LCC style capitalizes the first character of each word except the first word. For example, currencyCode="EUR".

53. Regarding the enumeration of values or code list text, it SHOULD be short but semantically sufficient and in English when there is no standard code list. The values and codes SHOULD be drawn from the common industrial property business language.

54. A limit of 35 characters for a name is recommended. When the same word is repeated in an element name, the second or following occurrences SHOULD be removed.

55. Element, attribute and type names MUST NOT include periods (.), spaces or other separators, or characters not allowed by W3C XML 1.0 for XML names except as specified in this Standard. For example, Office or domain prefixes (XX_UCC with XX in ST.3 code).

56. The characters used in enumeration value names described in this Standard are restricted to the following set: {a-z, A-Z, 0-9, period (.), comma (,), spaces, dash (-) and underscore (_)}.



HANDBOOK ON INDUSTRIAL PROPERTY INFORMATION AND DOCUMENTATION

Ref.: Standards - ST.66

page: 3.66.6

Acronyms and Abbreviations

57. XML element, attribute and type names **MUST NOT** use acronyms, abbreviations, or other word truncations, except as specified in this Standard or listed in Appendix D.
58. The acronyms and abbreviations listed in Appendix D **MUST** always be used instead of the complete extended name.
59. Acronyms and abbreviations at the beginning of an attribute declaration **MUST** appear in all lower case. All other acronym and abbreviation usage in an attribute declaration **MUST** appear in upper case.
60. Acronyms **MUST** appear in all upper case for all element declarations and type definitions.

XML Schema File naming Rules

61. These conventions will ensure that objects will be stored in a manner that will ensure consistency, uniformity, and comprehensiveness, and will be suitable for all aspects of storage and reuse.
62. Schema and style sheet filenames are recommended to follow a six-part naming rule. The six-part naming rule is illustrated below:

Part	Description	Syntax
[Office]	Use in case of office-specific artifact or union of several offices' codes.	AA (ST.3 code). It is to be omitted for an office generic version. For organizations not identified in WIPO Standard ST.3 or companies, they SHOULD not use two letter-codes but rather composed of three or four UCCs.
-	Delimiter.	A single dash.
[Domain]	Indication of the domain.	Aaa (variable length name, maximum length of 8 characters). Example: CTM = Community Trade Mark RCD = Registered Community Design
-	Delimiter.	A single dash.
Message or Service Name		Message or Service Name in UCC Example: CTMDownload ClassTerm
-	Delimiter.	A single dash.
Version	There are two options: 1. Version and sub-version (separated from the previous part by a dash). 2. Version by date (e.g. cyy-mm-dd).	
.	Delimiter.	A single period.
Extension	File extension (separated from the previous part by a dot).	Aaa (two to four characters). Example: xsd, xml, xsl



HANDBOOK ON INDUSTRIAL PROPERTY INFORMATION AND DOCUMENTATION

Ref.: Standards - ST.66

page: 3.66.7

Example: EM-CTM-Keyin-V1-0.xsd

[Office]	[Domain]	Message Service	Version	Extension
EM-	CTM-	Keyin-	V1-O	.xsd
		CTMDownload-	V3-2	.xsd
		CTMDownload	2007-02-25	.xml

Note: Field in square brackets [] is OPTIONAL.

63. File names SHOULD follow the above-mentioned tag naming rules. However a mapping can be defined locally if the rules cannot be applied due to technical constraints. Such local rules MUST be well defined and published for all potential users.

64. Schema file names SHOULD have their versions changed when an included modular schema is updated. With regard to the latest version of the XML schema, offices MAY provide two types of the latest version on their web site: a versioned schema which has an appropriated version number (e.g., WIPOST3Code-V2005-05-21.xsd, st36for66-V1-0.xsd) and a non-versioned schema without version number (e.g., WIPOST3Code.xsd, st36for66.xsd), which is a copy of (or refers to) the latest version of the XML schema. Other offices can refer to the latest version by indicating either the versioned schema with the latest version number or the non-version schema which is a copy of the latest version to avoid changing office's implemented codes whenever the schema is updated.

Miscellaneous XSD Rules

65. Restriction on field length MUST not be defined for the ST.66 XML Schema, but MAY be done for implementation schemas.

66. The <any> element SHOULD be used to offer extension and to keep the ST.66 XML Schema (Appendix B) opened to additional elements. It MUST not be used in implementation schema.

67. Elements SHOULD be declared with occurrence indicators. The occurrence indicators should not be declared explicitly when the required value is the default value.

68. The content or value within tags and attributes may be in any language, except enumerations.

69. A revision history of schema SHOULD not be inserted in the schema itself. The reference to the revision history, and the latest version number and date of the schema SHOULD be only given in the XML schema. The revision history SHOULD document changes with a date and description of each change, in reverse chronological order and published via the Office's web site. For example:

```
<xs:annotation>
  <xs:documentation>WIPO Standard ST.66 Model Schema Version 1.1, published in
  2007-12-03. The revision history is available on WIPO website at
  http://www.wipo.int/standards/en/xml_material/st66/revision-history/index.html
</xs:documentation>
</xs:annotation>
```

70. Contact points SHOULD be included in the prologue. For example:

```
<!-- Author : SDWG ST.66 Task Force -->
<!-- Contact : xml.standards@wipo.int -->
```

Naming Office-Specific Types and Elements

71. A namespace SHOULD be established for office-specific elements, where the office code (ST.3) becomes the prefix for identifying elements that are in that namespace.

72. Types that are not defined in the Standard can be defined as office-specific. Type names SHOULD have a prefix specific to the organization followed by a single dash. In the case of trademark offices, types SHOULD be prefixed by the two-letter office code as specified in WIPO Standard ST.3.



HANDBOOK ON INDUSTRIAL PROPERTY INFORMATION AND DOCUMENTATION

Ref.: Standards - ST.66

page: 3.66.8

73. For organizations/offices not identified in WIPO Standard ST.3 or companies, they SHOULD not use two-letter codes but rather codes composed of three or four uppercase letters.

74. As an alternative, a unique namespace SHOULD be established for office-specific elements, where the country code or company symbol becomes the prefix for identifying elements that are in that namespace.

External Entities

75. An external entity is any object that accompanies an XML document instance that is referenced from within the document instance. External entities are an integral part of a trademark document. Without them, the XML instance cannot be parsed, rendered, or understood successfully.

76. ~~XML document instance MIGHT be accompanied by external entities in base64 binary or refer to the filename of external entities given separately. The latter option SHOULD be recommended. (Check Annex F)~~
In the trademark domain, an external entity is most frequently an image, usually of the mark. External entities that are images of the mark SHALL conform to one or the other of the following profiles published by WIPO on its website. ⁽²⁾

JPEG

PNG

TIFF

GIF

77. Images can be embedded in an XML document instance as embedded binary images encoded in base64Binary that is the W3C XML Schema standard data type as well as references to external image files, i.e., external entities. However images SHOULD be referenced as external entities.

[The Appendices are available at: http://www.wipo.int/standards/en/xml_material/st66/]

[End of Standard]

⁽²⁾ At present, the SDWG Trademark Standards Task Force is preparing, for adoption as a WIPO Standard, a recommendation for the electronic management of the figurative elements of trademarks.