WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

**HANDBOOK ON INDUSTRIAL PROPERTY
INFORMATION AND DOCUMENTATION**

Ref.: Standards - ST.96

page: 3.96.0

STANDARD ST.96

Version 1.0

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

*Standard adopted by the Committee on WIPO Standards (CWS)
at its second session on May 4, 2012*

Editorial Note

Annexes V and VI of this Standard are drafts which require implementation and test by Offices to refine them as necessary. Subsequently, final proposal of Annexes V and VI will be prepared and submitted for consideration and adoption by the Committee on WIPO Standards.

TABLE OF CONTENTS

ANNEXES

The Annexes are available at: http://www.wipo.int/standards/en/xml_material/st96/

Annex I – XML Design Rules and Conventions
Annex II – IP Data Dictionary
Annex III – IP XML Schemas
Annex IV – XML Schema Technical Specification
Annex V – Schema Implementation Rules and Guidelines (draft)
Annex VI – Transformation Rules and Guidelines (draft)

WIPO STANDARD ST.96

Version 1.0

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

Standard adopted by the Committee on WIPO Standards (CWS)
at its second session on May 4, 2012

INTRODUCTION

1.   This Standard recommends the XML (eXtensible Markup Language) resources to be used for filing, publication, processing, and exchange of information for all types of industrial property (IP), i.e., patents, trademarks and industrial designs.

2.   This Standard recommends only the W3C XML schema language. For further information about the W3C (World Wide Web Consortium), see http://www.w3c.org/.

3.   This Standard comprises a set of recommendations, including rules and conventions for designing XML resources, and a common vocabulary for the IP community to exchange IP information in XML.

4.   This Standard is intended to:

(a)   enhance interoperability amongst industrial property offices (IPOs);

(b)   improve the harmonization among all types of IP information;

(c)   facilitate data compatibility, in particular, data transformability with WIPO Standards ST.36, ST.66 and ST.86;

(d)   improve the consistency of IP information; and

(e)   promote a common practice of implementation of XML schemas.

5.   This Standard refers to other WIPO Standards and industry standards including ISO Standards. Further details are available in the References Section below.

DEFINITIONS AND TERMINOLOGY

6.   The term "XML *resources*" is intended to refer to any of the components used to create and operate an XML implementation according to this Standard.

7.   The term "XML *schema*" is a language for describing the structure and constraining the contents of XML documents.

8.   The keywords MUST, MUST NOT, SHALL, SHOULD, SHOULD NOT, and MAY, when they appear in this Standard, are to be interpreted as described below. Non-capitalized forms of these words are used in the regular English sense. The definition of keywords is based on the Request for Comments 2119 - Key words for use in RFCs to Indicate Requirement Levels issued by the Internet Engineering Task Force.

(a)   MUST: This word, or the terms "REQUIRED" or "SHALL", means that the definition is an absolute requirement of the specification;

(b)   MUST NOT: This phrase, or the phrase "SHALL NOT", means that the definition is an absolute prohibition of the specification;

(c)   SHOULD: This word, or the adjective "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;

(d)   SHOULD NOT: This phrase, or the phrase "NOT RECOMMENDED", means that there may exist valid reasons in particular circumstances when the particular behavior is acceptable or even useful, but the full implications should be understood and the case carefully weighed before implementing any behavior described with this label; and

(e)   MAY: This word, or the adjective "OPTIONAL", means that an item is truly optional. An implementation that does not include a particular option MUST be prepared to interoperate with another implementation that does include the option, though perhaps with reduced functionality. In the same vein, an implementation that does include a particular option MUST be prepared to interoperate with another implementation that does not include the option (except, of course, for the feature the option provides).

9.   For the purposes of this Standard, the following definitions are given:

(a)   the expression "patent" includes such industrial property rights as patents for inventions, plant patents, design patents, inventors' certificates, utility certificates, utility models, patents of addition, inventors' certificates of addition and utility certificates of addition;

(b)   the expression "trademark" means trademark, service mark or another type of distinguishing representation of mark according to the definition of the mark in the legislation concerned, including but not limited to collective mark, certification mark or guarantee mark; and

(c)   the expression "industrial design" includes two-dimensional and three-dimensional features of shape and surface of objects, and thus covers both concepts of "designs" and "models" where a distinction is made between the former and the latter; the term "industrial designs" does not include design patents.

10.   *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.

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, deal with patent, trademark and/or industrial design documents.

12.   This Standard is intended to provide XML resources to be used for filing, publication, processing, and exchange of IP information.

13.   It is understood that this Standard cannot possibly include all XML schemas required by all IPOs. Instead, this Standard provides guidance for IPOs to implement office-specific schemas by using XML resources defined in this Standard.

14.   This Standard specifies:

(a)   the XML Design Rules and Conventions for Industrial Property (DRCs), i.e., Annex I, which provides a comprehensive set of design rules and conventions for the creation of XML schemas and instances used for all types of IP information.

(b)   the Data Dictionary, i.e., Annex II, which is data dictionary for IP information. It is intended to define data entities used for filing, processing, publication, and exchange of IP information;

(c)   the XML Schemas, i.e., Annex III, which is the set of XML schemas for components defined in Annex II. The XML schemas are intended to be used for inter-office communication and reused for office implementation;

(d)   the Schema Technical Specification, i.e., Annex IV, which provides technical documentation and details on the XML Schemas specified in Annex III;

(e)   the Schema Implementation Rules and Guidelines, i.e., Annex V, which provides guidance for customization of XML Schemas defined in Annex III to fit a specific business requirement of an IPO; and

(f)   the Transformation Rules and Guidelines, i.e., Annex VI, which provides guidance for transformation between XML instances conforming to this Standard and XML instances validating against ST.36, ST.66 or ST.86 .

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

HANDBOOK ON INDUSTRIAL PROPERTY
INFORMATION AND DOCUMENTATION

Ref.: Standards - ST.96

page: 3.96.3

REQUIREMENTS OF THE STANDARD

General

15.   The XML Design Rules and Conventions for Industrial Property (DRCs), i.e., Annex I, is the foundation of this Standard.

16.   Conformance is important in the context of data exchange. ST.96 conformance at the instance and schema level means that there are no constraint violations when validating an instance against an ST.96 Schema which is defined in Annex III. An ST.96 conformant XML instance is an instance that validates against an ST.96 Schema. An XML schema is an ST.96 conformant schema if it sustains constraints expressed by an ST.96 Schema. For data exchange, an XML document instance MUST be an ST.96 conformant XML instance.

17.   Implementation of XML schemas that is compatible to this Standard MUST be carried out according to the DRCs, Annex I, and SHOULD reuse XML schemas defined in Annex III following the Schema Implementation Rules and Guidelines, Annex V.

18.   Each dictionary entry name defined in the IP Data Dictionary, i.e., Annex II, MUST be unique. A dictionary entry name SHOULD be composed of object class, property and representation terms of which definition is given in Annex I to this Standard. Each term MAY consist of multiple words. Each word in a term is separated from the others by a space character ( ). Each word within a term must start with an uppercase letter. Each term in a dictionary entry name is separated from the others by dot character (.). The terms in the dictionary entry name are separated by a period (.) and a space character ( ).

19.   Maintaining compatibility with existing documents using WIPO Standards ST.36, ST.66 and ST.86 is one of the primary concerns for this Standard. Therefore, this Standard seeks the necessary degree of compatibility and convertibility with WIPO Standards ST.36, ST.66, and ST.86 in order to ensure that data can be processed satisfactorily for the business needs of IPOs and IP information suppliers. While an attempt has been made to incorporate improvements over ST.36, ST.66, and ST.86, not all national requirements were captured. Consequently, this Standard leaves the transformation of any remaining national elements as the responsibility of Offices which have extended ST.36, ST.66, and ST.86.

20.   In this Standard, many XML schemas use the model of choice between structured format and unstructured format at the same time. In all cases, structured text is preferred to unstructured text or images.

XML schema structure

21.   This Standard recommends a smaller modular set of XML schemas rather than a big monolithic XML schema. It aims to promote sharing and reusing individual elements or logical components defined as a group in a schema.

22.   In this Standard, XML schema modules are categorized into Common Component and specific IP type component, i.e., Patent Component, Trademark Component and Design Component schemas. Common Component SHOULD be context-neutral (or business independent) and shared by components of, at least, two IP types.

Identification of types, elements and attributes: Namespace

23.   Namespace MUST be used to identify Types, elements and attributes.

24.   This Standard defines the following four namespaces:

–   http://www.wipo.int/standards/XMLSchema/Common/ for Common Component Schemas;

–   http://www.wipo.int/standards/XMLSchema/Patent/ for Patent Component Schemas;

–   http://www.wipo.int/standards/XMLSchema/Trademark/ for Trademark Component Schemas; and

–   http://www.wipo.int/standards/XMLSchema/Design/ for Design Component Schemas.

Naming office-specific schemas

25.   A namespace SHOULD be established for office-specific components, i.e., Types, elements and attributes, where the two-letter office code as specified in WIPO Standard ST.3 becomes the prefix for identifying components that are in that namespace.

26.   Types, elements and attributes names SHOULD have a prefix specific to the IPO or company. In the case of IPOs, Type and element SHOULD be prefixed by the two-letter office code as specified in WIPO Standard ST.3, for example, `USMathType`. IPOs not identified in WIPO Standard ST.3 or companies MUST NOT use the ST.3 two-letter codes or prefix defined in this Standard, but rather codes composed of three or four uppercase letters. The prefix codes SHOULD be registered in the Registry of this Standard.

External entities

27.   An external entity is any object that accompanies an XML document instance and contains a reference from the XML document instance. External entities may be associated with patent documents, trademark documents or design documents. The XML instance cannot be parsed, rendered, or understood successfully when the referenced external entities do not exist. In the case of patent documents, external entities are most frequently pages of drawing, but could also include embedded images, computer software listings, mathematical formulas, chemical formulas, tables, sequence listings, undefined characters, or character entities.

28.   An external entity is most frequently an image. This image is usually an image of drawings in patent documents, of the figurative element of marks in trademarks, and of the representation of industrial designs in industrial design documents. External entities that are images SHOULD conform to one of the following image formats:

-      JPEG, PNG, TIFF or GIF for patent documents;

-      Image formats recommended in WIPO Standard ST.67 for trademarks; and

-      JPEG, PNG, TIFF or GIF for industrial designs.

Industry-standard schemas

29.   Where appropriate to the content of a document, that is, where the content is not unique to the industrial property domain, industry-standard schemas SHOULD be used. The approved Industry-standard schemas SHOULD be stored in the Repository specified by this Standard.

30.   The following industry-standard schemas are referred to in this Standard:

–   OASIS XML table schema version 1.0 available at:
        http://www.oasis-open.org/docbook/xmlschema/1.0b1/;

–   MathML3 available at: http://www.w3.org/TR/MathML3/;

–   ISO 3166-1 Country Code schema;

–   ISO 639-1 Language Code schema; and

–   ISO 4217 Currency Code schema.

REFERENCES

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

        (a)   WIPO Standard ST.3 – Two-Letter Codes for the Representation of States, Other Entities and Intergovernmental Organizations;

        (b)   WIPO Standard ST.25 – Presentation of nucleotide and amino acid sequence listings

        (c)   WIPO Standard ST.36 – Processing of Patent Information Using XML;

        (d)   WIPO Standard ST.66 – Processing of trademark information using XML

        (e)   WIPO Standard ST.67 – Electronic management of the figurative elements of trademarks

        (f)   WIPO Standard ST.80 – Bibliographic data relating to industrial designs

        (g)   WIPO Standard ST.86 – Processing of industrial design information using XML

        (h)   W3CXML Schema Part 1: Structures (http://www.w3.org/TR/xmlschema-1/) and XML Schema, Part 2: Datatypes (http://www.w3.org/TR/xmlschema-2/).

(i)   ISO/IEC 11179-5 Information technology – Metadata registries (MDR) – Part 5: Naming and identification principles;

(j)   ISO 3166-1 – Codes for the representation of names of countries and their subdivisions – Country Codes;

(k)   ISO 639-1 – Codes for the representation of names of languages – Part 1: Alpha2-code;

(l)   ISO 4217 – Codes for the representation of currencies and funds;

(m)   ISO/IEC 10646 – Information technology – Universal Multiple-Octet Coded Character Set (UCS);

(n)   Internet Engineering Task Force (IETF) Request For Comments (RFC) 2119 – Key words for use in RFCs to Indicate Requirement Levels.

[The Annexes are available at: http://www.wipo.int/standards/en/xml_material/st96/]

[End of Standard]