STANDARD ST.96

Version 76.0

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

Revision approved by the XML4IP Task Force of the Committee of WIPO Standards (CWS) on ~~October~~ April 3, 2023~~2~~

Table of Contents

~~STANDARD ST.96 ..... 1~~
~~WIPO STANDARD ST.96 ..... 2~~
INTRODUCTION 2
DEFINITIONS AND TERMINOLOGY 2
SCOPE OF THE STANDARD 3
REQUIREMENTS OF THE STANDARD 4
General 4
XML schema structure 4
Identification of types, elements and attributes: Namespace 4
Naming office-specific schemas 5
External entities 5
Industry-standard schemas 5
REFERENCES 5

ANNEXES

The Annexes are available at: ~~http://www.wipo.int/standards/en/st96/v6-0/~~http://www.wipo.int/standards/en/st96/v7-0/

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
Annex VI – Transformation Rules and Guidelines
Annex VII – Example XML Instances

WIPO STANDARD ST.96

Version 76.0

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

Revision approved by the XML4IP Task Force of the Committee on WIPO Standards (CWS) on ~~October~~ April 3, 2023~~2~~

INTRODUCTION

1. This Standard recommends the XML (eXtensible Markup Language) resources to be used for filing, publication, processing, and exchange of information for ~~all~~ the followingtypes of intellectual property (IP)~~,~~ ~~i.e.,~~ patents, trademarks, industrial designs, geographical indications and copyright. This Standard can be extended to other IP types such as Plant Breeder's Rights (PBR)[1].

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 intellectual 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;

[1] The International Union for the Protection of New Varieties of Plants (UPOV) implements ST.96 in relation to Plant Breeder's Rights, i.e., PVP-XML. Further information on PVP-XML is available on UPOV's website at: https://www.upov.int/upovprisma/en/pvpxml.html.

(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;

(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;

(d) the expression “geographical indication” means a sign used on products that have a specific geographical origin and possess qualities or a reputation that are due to that origin; and

(e) the expression “copyright’ includes rights that creators have over their literary and artistic works and the expression “copyright orphan work” includes works that are protected by copyright, but the author cannot be identified or found.

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 intellectual property laws or international intellectual property conventions, deal with data and documents of patent, trademark, industrial design, geographical indication and/or copyright orphan work.

12. This Standard is intended to provide XML resources to be used for filing, publication, processing, and exchange of IP data and 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 and data exchange among IPOs by using XML resources defined in this Standard.

14. This Standard specifies:

(a) the XML Design Rules and Conventions (DRCs) for IP data, 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 a data dictionary for IP information. It is intended to define data entities used for filing, processing, publication, and exchange of IP data and 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;

(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 validated against WIPO Standards ST.36, ST.66 or ST.86; and

(g) the Example XML Instances, i.e., Annex VII, which provides example XML instances.

REQUIREMENTS OF THE STANDARD

General

15. The XML DRCs, i.e., Annex I, is the foundation of this Standard.

16. All elements, attributes and enumeration values of this Standard MUST be listed in the IP Data Dictionary, i.e., ~~Annex~~ Annex II.

17. 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.

18. Every version of Annex III MUST be released in its entirety including updated and unmodified XML Schema components.

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

20. 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.

21. 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

22. This Standard recommends modular set of XML schemas at component level during design and flattened XML schemas at document level for implementation. In design mode, each element, attribute, and type is defined in its own schema file. They are used as building blocks to promote sharing and reusing through `xsd:include` for the components in the same namespace and `xsd:import` for components from different namespaces. Flattened schema at document level resolves all the `xsd:~~include~~ include` by copying all the components belonging to each namespace into one schema file and imports documents schemas from other namespaces using `xsd:import` statement. The flattened approach provides the efficiency and convenience in implementation. The flattened schemas should be provided with every release of the design schemas.

23. 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

24. Namespace MUST be used to identify types, elements and attributes.

25. This Standard defines namespaces for Common Components, Patent Components, Trademark Components and Design Components as listed in Annex I.

Naming office-specific schemas

26. A namespace SHOULD be established for office-specific components, i.e., types, elements and attributes, where the two-letter office code in lower case as specified in WIPO Standard ST.3 becomes the prefix for identifying components that are in that namespace, for example, "`uspat`", "`ustmk`".

27. New or modified types, elements and attributes names SHOULD be defined in IPO namespace and/or have a prefix in the component name specific to the IPO or company. Namespace prefix SHOULD be a combination of the two-letter office code in lower case as specified in WIPO Standard ST.3 and one of namespace prefix defined in this Standard, i.e., `com, pat, tmk, cpy, gin` and `dgn`. For example, `uspat:MathType`.

External entities

28. 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 2D or 3D images[2], multimedia files, sound recordings, 3D models[3], computer software listings, mathematical formulas, chemical formulas, tables, sequence listings, undefined characters, or character entities.

29. 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 or 3D models SHOULD conform to one of the following ~~image~~ formats:

- JPEG, PNG, TIFF, GIF or SVG for patent documents;
- Image formats recommended in WIPO Standard ST.67 for trademarks and geographical indications;
- 3D image or 3D model formats recommended in WIPO Standard ST.91 for patents, trademarks, industrial designs and integrated circuit topography;
- JPEG, PNG, TIFF, GIF or SVG for industrial designs; and
- JPEG, PNG, TIFF, GIF or SVG for copyright orphan works.

30. An external entity may also be a multimedia or sound file. External entities that are multimedia files should comply with one of the following formats:

- MP3 or WAV for a sound marks as recommended in WIPO Standard ST.68; and
- MP4 or WebM for a multimedia file as recommended in WIPO Standard ST.69.

Industry-standard schemas

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

32. 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/; and
- MathML3 available at: http://www.w3.org/TR/MathML3/.

REFERENCES

33. The following Standards and resources are of relevance to this Standard:

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

[2] 3D Images: Digital images that represent objects displayed in three dimensions such as 3D photos and stereoscopy (See paragraph 3 of WIPO Standard ST.91)

[3] 3D model: An electronic file that is created by specialized software, for mathematically representing the surface of an object's visual representation in three dimension (See paragraph 3 of WIPO Standard ST.91)

WIPO Standard ST.9	Bibliographic data on and relating to Patents and SPCs;
WIPO Standard ST.16	Identification of different kinds of patent documents;
WIPO Standard ST.25	Presentation of nucleotide and amino acid sequence listings;
WIPO Standard ST.26	Presentation of nucleotide and amino acid sequence listings using XML;
WIPO Standard ST.27	Exchange of Patent Legal Status Data;
WIPO Standard ST.36	Processing of Patent Information Using XML;
WIPO Standard ST.37	Authority File of Published Patent Documents;
WIPO Standard ST.60	Bibliographic data relating to marks;
WIPO Standard ST.61	Exchange of Trademark Legal Status Data;
WIPO Standard ST.66	Processing of trademark information using XML;
WIPO Standard ST.67	Electronic management of the figurative elements of trademarks;
WIPO Standard ST.68	Electronic Management of sound marks;
WIPO Standard ST.69	Electronic management of motion and multimedia marks;
WIPO Standard ST.80	Bibliographic data relating to industrial designs;
WIPO Standard ST.86	Processing of industrial design information using XML;
WIPO Standard ST.87	Exchange of industrial design legal status data;
WIPO Standard ST.90	Processing and communicating IP data using web APIs;
WIPO Standard ST.91	Digital Three-Dimensional (3D) models and 3D Images;
WIPO Standard ST.97	Processing of intellectual property data using JSON;
W3CXML Schema	XML Schema Part 1: Structures (http://www.w3.org/TR/xmlschema-1/) and XML Schema, Part 2: Datatypes (http://www.w3.org/TR/xmlschema-2/);
International Standard ISO/IEC 11179-5	"Information technology – Metadata registries (MDR) – Part 5: Naming and identification principles";
International Standard ISO 3166-1	"Codes for the representation of names of countries and their subdivisions – Country Codes";
International Standard ISO 639-1	"Codes for the representation of names of languages – Part 1: Alpha2-code";
International Standard ISO 4217	"Codes for the representation of currencies and funds";
International Standard ISO/IEC 10646	"Information technology – Universal Multiple-Octet Coded Character Set (UCS)";
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/st96/v6-0/~~http://www.wipo.int/standards/en/st96/v7-0/]

[Annexes follow]