



Australian Government

IP Australia

Application Development Standards

***XML GUIDELINES FOR
IP AUSTRALIA DEVELOPERS***

Version: 0.4

Date: 12/02/2007

Status: Published

Author: Samantha Hoy

Documentation and Approvals

Document Location

The master copy of this document is held at the following location:

<http://intranet/intranet/Forum4/Topic613/IPA XML Guidelines.doc>

Revision History

This document is under version control

Revision Date	Version No.	Author	Description	Changes Marked
	0.1		First draft	N
31/01/2007	0.2		Comments and examples provided by Stephen Davies and Andy Cheng	N
09/02/2007	0.3		<ul style="list-style-type: none">• Changed reference for XML version 1.1 to 1.0 given common problems with parsing version 1.1.• Changed version extension format in file names to use dash instead of underscore; WIPO and other IP Offices use this format, eg v1-0 instead of v1_0• Added new section "Other Exceptions to the Rules" given current SAP limitation.	N
12/02/2007	0.4	Steve Luckie	Rewritten to suit the Corporate document template	N

Document Approval

Name	Signature	Title	Date	Version
Dan Collins		Director, Process and Systems Delivery	27/02/2007	0.4

Table of Contents

1	General Introduction.....	4
1.1	Subject	4
1.2	Introduction	4
1.2.1	Bibliographic documents.....	4
2	Definitions	6
2.1	Acronyms and terms	6
3	Registry for Published Schema.....	7
4	Namespaces.....	8
4.1	Naming and versioning	8
4.2	Namespace declarations	8
5	Naming of Schema and Stylesheet File Names	10
6	Document Content.....	11
7	Naming of Elements, Attributes and Types.....	12
8	Other Exceptions to the Rules.....	14

1.1 Subject

Naming and versioning convention of XML related components within IP Australia.

1.2 Introduction

This document will provide guidance to IP Australia developers during the creation of new schemas and stylesheets, and related element, type and attribute contents. The intent of these guidelines is to follow WIPO documented standards where applicable. Consideration for proposed WIPO standards was anticipated whilst incorporating generally accepted XML naming conventions for the purposes of standardising all future XML artefacts.

These guidelines have been developed with a view to:

1. Address consistency around versioning and naming of XML related artefacts;
2. Minimise binding complications intrinsic in older naming conventions that include the use of dashes (-) in the naming of elements and attributes;
3. Distinguish when, and when not, to incorporate International Comment Element (ICE) artefacts; and
4. Comply with W3C XML Schema Language standards.

The result of these guidelines will be an XML environment that contains a mix of ST.36 related ICE naming conventions and more current generic XML naming conventions based on W3C and ebXML standards.

It should be noted that this position is currently being incorporated into the future ST.66 standard. ICE elements common to both Patent and Trademarks will be incorporated into the ST.66 dictionary using ST.36 naming conventions.

1.2.1 Bibliographic documents

The following is a list of the reference documents and standards referred to or used as a source of background information during the production of this document:

1. WIPO Standard ST.36 Recommendation for the Processing of Patent Documents Using XML Sep 2005. (<http://www.wipo.int/scit/en/standards/pdf/03-36-01.pdf>).
2. WIPO International Comment Elements (ICE). (<http://www.wipo.int/scit/en/standards/st36-xml-dtd/epct-element-encyclopedia-20050220.xls>).
3. WIPO Standard ST.66 Recommendation for the Processing of Trademark Documents Using XML v.0.3 draft Dec 2006.

Note: Please contact the IPA International Information Management Group (IIMG) for further information and directions to the most up-to-date version of this document, as it is currently under review.

4. WIPO Standard ST.3: Recommended Standard on Two-Letter Codes for the Representation of States, Other Entities and Intergovernmental Organizations. (<http://www.wipo.int/scit/en/standards/pdf/03-03-01.pdf>)

5. ebXML Technical Architecture Specification v1.0.4, ebXML Technical Architecture Project Team. (<http://www.ebxml.org/specs/ebTA.doc>)
6. ISO/IEC 10646 – Universal Character Set (UCS) – Unicode 3.2 UTF-8. (See short version <http://www.nada.kth.se/i18n/ucs/unicode-iso10646-overview.html>, or http://en.wikipedia.org/wiki/Universal_Character_Set)
7. W3C Namespaces in XML 1.0 (second edition). <http://www.w3.org/TR/2006/REC-xml-names-20060816/>

2.1 Acronyms and terms

Term	Definition
ICE	International Comment Elements
WSDL	Web Services Description Language

3 REGISTRY FOR PUBLISHED SCHEMA

(See XML Infrastructure document - yet to be written)

4.1 Naming and versioning

IP Australia namespaces should include the following:

1. Public Address: ie, <http://www.ipaustralia.gov.au>;
2. Business line: eg, patents, trademark, designs, etc;
3. Document type: eg, schema, wsdl, stylesheet, etc;
4. Schema name: eg, pams-application-lodgement;
5. Year: eg, 2005, 2007; and
6. Version number: 1, 2, 3, etc. (unique within year, always commencing at 1)

For example:

<http://www.ipaustralia.gov.au/corporate/stylesheet/financial-invoice/2005/1>; or
<http://www.ipaustralia.gov.au/patents/schema/pams-application-lodgment/2007/1>

4.2 Namespace declarations

1. Every schema module, MUST have a namespace declared using the `xsd:targetNamespace` attribute.
2. All XML schemas MUST declare the W3C schema namespace.
3. Every defined or used schema MUST have its own unique namespace.
4. Published namespaces MUST never be changed. Modifications need to have their own new namespace. This is a WIPO requirement and it also avoids problems where consumers may cache schema documents and also prevents versioning issues.
(See *ST.66.ver0.3_final.doc - paragraph 402*).
5. There SHALL be no default namespaces. That is, for example, both `XMLSchema` and `targetNamespace` MUST be explicitly qualified. This approach, even though quite cluttered, is more consistent for all types of schema.
(See *ST.66.ver0.3_final.doc - paragraph 41*).

```
<?xml version="1.0" encoding="UTF-8" ?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  targetNamespace="http://www.ipaustralia.gov.au/corporate/schema/example/2007/1"
  xmlns:lib="http://www.wipo.int"
  xmlns:tns="http://www.ipaustralia.gov.au/corporate/schema/example/2007/1"
  elementFormDefault="qualified">
<xsd:include schemalocation="xxx.xsd"/>
<xsd:import namespace="http://www.wipo.int" schemaLocation="../../schemaLocation.xsd"/>
</xsd:schema>
```

Note: The exception to this rule is when ICE Elements are used since these elements are currently not defined with a target namespace.

Example of how to include ICE elements:

```
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema"
targetNamespace="http://www.ipaustralia.gov.au/corporate/schema/example/2007/1"
  xmlns:tns="http://www.ipaustralia.gov.au/corporate/schema/example/2007/1"
  elementFormDefault="qualified">
<xsd:include schemalocation="invention-title.xsd"/> ( This is a reference to an ICE
element)
</xsd:schema>
```

6. External schema references SHOULD use the "Include" construct. The including and included schemas MUST have the same target namespace. (See *ST.66.ver0.3_final.doc - paragraph 42*).
7. External schema references with a different namespace MUST use the 'import' construct.

5 NAMING OF SCHEMA AND STYLESHEET FILE NAMES

The naming convention for the file name of schemas and stylesheets should follow a convention where separate words describing the file name should be delimited with dashes.

All documents must be versioned and the version number should be distinguished using a lower case “v” followed by the version number. Any sub-versions should be depicted using a dash and a sub-version number.

The following table explains this convention.

Part	Description	Syntax
Process	Description of process or data being addressed	Eg: national-phase-request
-	Separator	A single dash
Version	Version and sub-versions separated by dashes.	v1-1
.	Separator	A single period (full stop)
Extension	File extension	Eg: xsd, xsl, wsd

For example:

national-phase-request-v1-1.xsd

6 DOCUMENT CONTENT

Document instances **MUST** include the following processing instruction as the first line in the file.

Note: Only UTF-8 is supported in this Standard.

Although XML defaults to UTF8, WIPO would like us to make it explicit.

```
<?xml version="1.0" encoding="UTF-8" ?>
```

7 NAMING OF ELEMENTS, ATTRIBUTES AND TYPES

1. When creating new documents, always check the International Common Elements (ICE) documentation before creating new elements, type or attributes as all outgoing patent related xml must currently comply with ST.36 and use the existing ICE element.
2. When changing the content of a documented ICE element for the purposes of international/external publication, ST.36 rules for naming of Office specific elements apply. In short, the rules are as follows:
 - a. Each Office-specific element name shall be preceded by the ST.3 country code of the office that owns the element. In our case this would be au.
 - b. The code shall be separated from the element name by either a hyphen (-) or a colon (:). For example, jp:fterm, ep-printer-name. The colon is used only where the owning office is implementing W3C XML namespaces.
 - c. All element names should be common words from the English Language.
 - d. Where more than one word is required for an element name, the words should be separated by a hyphen (-), NOT an underscore.
 - e. Element names should be limited to the following set of characters - abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789-
 - f. Names should be descriptive, not mnemonic or abbreviated, as far as practical

(See WIPO ST.36 03-36-01.pdf, paragraphs 19-36)

3. When creating new documents for internal use, IP Australia namespaces must be used
4. For all new element attribute and types created, the following standards rules should be followed:
 - a. Names should be descriptive, not mnemonic or abbreviated, as far as practical.
 - b. Acronyms SHOULD be avoided, but in cases where they are used, the capitalization SHALL remain (example: XMLSignature).
 - c. Underscore (_), periods (.) and dashes (-) MUST NOT be used (don't use: header.manifest, stock_quote_5, commercial-transaction, use HeaderManifest, stockQuote5, CommercialTransaction instead). When including patent elements specified in the ICE documentation, dashes are an exception to this rule.
 - d. Element names MUST be in upper camel case (UCC). UCC style capitalizes the first character of each word which compounds the name.

For example:

```
<element name="AddressCountryCode" type="xs:string"/>
```

- e. Type names MUST be in UCC + Suffix Type.

For example: LanguageCodeType.

```
<complexType name="LanguageCodeType"> . . . </complexType>
```

- f. Attribute names MUST be in lower camel case (LCC). LCC style capitalizes the first character of each word except the first word.

For example:

```
<attribute name="currencyCode" type="xs:string"/>
```

(See *ebXML Technical Architecture Specification v1.0.4, section 4.3*)

8 OTHER EXCEPTIONS TO THE RULES

These exceptions are provided for instances where IP Australia's current environment prevents the implementation of the guidelines discussed in this document. They should only be used in circumstances where there is no workaround. It is hope that these uses will be kept to a minimum.

1. Within IP Australia's current SAP environment, there is an existing limitation on the length of namespace that can be processed. SAP XI cannot handle namespaces longer than 60 characters. For this reason, the extension of the version information specified in these guidelines is not mandatory for xml that will need to be processed through the SAP environment. Once this limitation is addressed in the SAP environment, versioning on namespaces should be incorporated.