

# TABLE OF CONTENTS

---

## TABLE DES MATIERES

*In English:*

<b>HS STRUCTURED FILES .....</b>	<b>pages 2 – 13</b>
<b>GENERAL NAMING CONVENTIONS OF THE STRUCTURED FILES .....</b>	<b>pages 2– 3</b>
<b>HS NOMENCLATURE FILES (CSV, XML, TXT) .....</b>	<b>pages 4 – 7</b>
<b>HS ALPHABETICAL INDEX FILES .....</b>	<b>page 8</b>
<b>HS EXPLANATORY NOTES FILES .....</b>	<b>pages 9 – 10</b>
<b>HS CLASSIFICATION OPINIONS FILES .....</b>	<b>pages 11 – 12</b>

*En Français :*

<b>FICHIERS STRUCTURES DU SH .....</b>	<b>pages 13 – 23</b>
<b>CONVENTIONS DE NOMMAGE GENERALES DES FICHIERS STRUCTURES .....</b>	<b>pages 13 – 14</b>
<b>FICHIERS DE LA NOMENCLATURE DU SH (CSV, XML, TXT) .....</b>	<b>pages 15 – 18</b>
<b>FICHIERS DE L'INDEX ALPHABETIQUE DU SH .....</b>	<b>page 19</b>
<b>FICHIERS DES NOTES EXPLICATIVES DU SH .....</b>	<b>pages 20 – 21</b>
<b>FICHIERS DES AVIS DE CLASSEMENT DU SH .....</b>	<b>pages 22 – 23</b>

# HS STRUCTURED FILES

## I- GENERAL NAMING CONVENTIONS OF THE STRUCTURED FILES

The WCO uses a naming convention for the distribution of files from these publications. The filename (excluding extensions) respects this convention.

Filenames are made up of several characteristics separated by a dash ("-"). The size of the name varies according to the characteristics present, which are specific to the file.

The characteristics are as follows (compulsory characteristics are in bold):

<b>Characteristic</b>	<b>Description</b>	<b>Possible values</b>
<b>Publisher name</b>	Comprises the name of the publisher of the publication from which the file is taken	Always "wcoomd"
<b>Collection</b>	Comprises the abbreviation of the collection to which the publication belongs, from which the file is taken	HS: Harmonized System
<b>Publication</b>	Comprises the abbreviation of the publication from which the file is taken	NAME: Nomenclature EN: Explanatory Notes CCO: Compendium of Classification Opinions COM: Goods database AI: Alphabetic Index
<b>Main version</b>	Comprises the prefix "V" and the main revision of the publication. For instance, for the harmonized system, the version commonly called 2007 is the 4th revision	V4: Harmonized System 2007 V5: Harmonized System 2012 V6: Harmonized System 2017
<b>Version date</b>	This is the "year" version of the publication, in the format "YYYY".	2007, 2012, 2017
Subordinate version	Comprises the prefix "e" (for erratum) and the erratum number. Only used in terms of the nomenclature of the harmonized system.	All values are possible in theory, with the specific case "e0" corresponding to the original version of the main version
Revision number	Comprises the prefix "as" (for "amending supplement") then the supplement number. Only used in terms of Explanatory Notes and Compendium of Classification Opinions of the Harmonized System, or the term "full" if the version is a full one.	All values are possible in theory. The absence of this characteristic corresponds to the original version of the main version.
<b>Adoption date</b>	Comprises the adoption date in the form "YYYYMM" (4-figure year, 2-figure month) by WCO members of the publication	All values are possible in theory. In practice, YYYY starts from 2007 and MM is obviously a number between "01" for January and "12" for December
<b>Format</b>	Comprises the file format	CSV: Delimited text TXT: Fixed-size text DOC: Word files PDF: PDF files XML: XML files XHTML: XHTML files
<b>Language(s)</b>	Comprises the languages present in the file (each language is separated by the "_" character).	En: English Es: Spanish Fr: French It: Italian Ru: Russian
Simulation	The "simulation" suffix can be used to define test files.	

## Actual examples of compliant file names:

### ***wcoomd-sh-nom-v5-2012-e0-201006-csv-en\_fr.zip***

ZIP file containing the original version of the 2012 nomenclature from June 2010 in CSV format in English and French.

### ***wcoomd-sh-nom-v4-2007-e3-200812-xml-en\_es\_fr\_it\_ru.zip***

ZIP file containing the 3rd erratum with 2007 nomenclature from December 2008 in XML format in English, Spanish, French, Italian and Russian.

### ***wcoomd-sh-ne-v5-2012-full-201006-pdf\_fr.zip***

ZIP file containing all explanatory notes of the 2012 version of the harmonized system from June 2010 in PDF format in French.

### ***wcoomd-sh-ne-v5-2012-as1-201112-doc\_en-simulation.zip***

ZIP file containing only the explanatory notes amended by AS1 of the 2012 version of the harmonized system in English from December 2011 in Word format, for simulation purposes only (unofficial data).

## II- HS NOMENCLATURE FILES

### ASCII delimited files (CSV)

#### Naming conventions

Updated versions available:

**wcoomd-sh-nom-v4-2007-e3-200812-csv-en\_fr.zip**

**wcoomd-sh-nom-v5-2012-e0-201006-csv-en\_fr.zip**

**wcoomd-sh-nom-v6-2017-e0-201606-csv-en\_fr.zip**

#### File description

The tab character ([TAB] CHR(9)) is used as a delimiter.

The first line contains the column headings.

The file is encoded in UTF-8.

Position	Name	Type	Description
1	Level	Numerical	Contains the level of the entry in the classification 1: Section 2: Chapter (HS2) 3: Position (HS4) 4: Sub-Position (HS6) or Grouping Sub-Position (HS5)
2	Father	Alphanumerical	Contains the connector line code in the form: <b>9999</b> (four Arabic numeral prefix): only for sections, specifies to which version of the nomenclature the section is connected (for instance "2012") <b>9999.X</b> (suffix in the form of a Roman numeral): only for chapters, specifies to which section the chapter is connected <b>9999.99</b> (suffix in the form of two Arabic numerals): only for positions, specifies to which chapter the position is connected <b>9999.9999</b> (suffix in the form of four Arabic numerals): for 6-figure sub-positions ending in "0" or for grouping sub-positions. Specifies the grouping position to which the sub-position is connected. <b>9999.99999</b> (suffix in the form of five Arabic numerals): only for 6-figure sub-positions ending in a figure from "1" to "9". Specifies the grouping sub-position to which the sub-position is connected
3	Hs_code	Alphanumerical	Contains the line's HS code in the form: <b>9999.X</b> (suffix in the form of a Roman numeral): only for sections <b>9999.99</b> (suffix in the form of two Arabic numerals): only for chapters (HS2) <b>9999.9999</b> (suffix in the form of four Arabic numerals): for positions (HS4). <b>9999.99999</b> (suffix in the form of five Arabic numerals): for grouping sub-positions (HS5) <b>9999.999999</b> (suffix in the form of six Arabic numerals): for sub-positions (HS6)
4	LANG1	Alphanumerical	Line heading, in the language specified in the first line
5	LANG2	Alphanumerical	Line heading, in the language specified in the first line
...			

The number of columns varies according to the number of languages present in the file. The number of lines is constant for a given version of the nomenclature, regardless of the number of languages present (for instance, the original 2012 nomenclature file contains 7,504 lines).

The end of each line is signalled by a carriage return.

## XML files

### Naming conventions

Updated versions available:

**wcoomd-sh-nom-v4-2007-e3-200812-xml-en\_fr.zip**

**wcoomd-sh-nom-v5-2012-e0-201006-xml-en\_fr.zip**

**wcoomd-sh-nom-v6-2017-e0-201606-xml-en\_fr.zip**

### File description

This file complies with the XSD diagram available on the WCO publications website.

URI: [http://wcoomdpublications.org/schemas/HS\\_converted.xsd](http://wcoomdpublications.org/schemas/HS_converted.xsd)

The file is encoded in UTF-8. It contains as many "NOMENCLATURE" nodes as there are lines, and for each of them, as many "LABEL" nodes as there are languages.

Each "LABEL" node receives the line heading in the language indicated in the "lang" attribute. The languages comply with ISO 639-1 codification (alpha-2):

en: English  
es: Spanish  
fr: French  
it: Italian  
ru: Russian

### Further information regarding "NOMENCLATURE" node attributes

The <NOMENCLATURE> node comprises of several attributes, some of which are optional (see XSD diagram). The table below specifies certain characteristics to make it easier to understand the XML file. The attributes shown in bold are compulsory.

Attribute	Description	Possible values
<b>Year</b>	Contains the main HS version (year) to which the line refers. The year is encoded numerically over 4 characters (9999).	2007 2012
<b>Section</b>	Contains the section number, in Roman numerals, to which the line is connected.	"I" to "XXI"
<b>Order</b>	Contains the logical order number in the classification that only contains numerals. Each level is represented by 2 digits, except for grouping sub-positions. Sections have 2 digits. Chapters have 4 digits. Positions have 6 digits. Sub-positions have 8 digits. Grouping sub-positions have 7 digits.	<b>Digits 1 and 2:</b> the section number in Arabic numerals, from 01 to 21 <b>Digits 3 and 4:</b> the chapter number from 01 to 97 <b>Digits 5 and 6:</b> the position number (variable) <b>Digits 7 and 8:</b> the sub-position number (variable) When the number 8 alone is missing, this represents a grouping sub-position
<b>Level</b>	Contains the level of the entry in the classification 1: Section 2: Chapter (HS2) 3: Position (HS4) 4: Sub-Position (HS6) or Grouping Sub-Position (HS5)	
Chapter	Contains the chapter number over 2 digits (equivalent to digits 3 and 4 of the "order" attribute)	<b>01 to 97</b>
Heading	Contains the position number in the chapter over 2 digits (equivalent to digits 5 and 6 of the "order" attribute)	<b>Variable</b>
Subheading	Contains the sub-position number in the position	<b>variable</b>

## Fixed-size ASCII files (TXT)

### Naming conventions

Updated versions available:

***wcoomd-sh-nom-v4-2007-e3-200812-txt-en\_fr.zip***

***wcoomd-sh-nom-v5-2012-e0-201006-txt-en\_fr.zip***

***wcoomd-sh-nom-v6-2017-e0-201606-txt-en\_fr.zip***

### File description

Each line has a fixed size of 1,008 characters. The line ends with a carriage return.

The file is encoded in UTF-8.

The first line contains the adoption date of the version provided.

Position	Length	Type	Description
1	1	Numerical	Contains the level of the entry in the classification 1: Section 2: Chapter (HS2) 3: Position (HS4) 4: Sub-Position (HS6) or Grouping Sub-Position (HS5)
2	6	Alphanumerical	<b>HS Code</b> <b>Section:</b> the code is provided in numeric form (and must be translated by you into Roman numerals) over two digits, completed with blanks ON THE LEFT. <b>Chapter:</b> the code is provided over two digits, completed by blanks ON THE LEFT <b>Position:</b> the HS4 code is provided over four digits, completed by blanks ON THE LEFT Sub-Position / Grouping Sub-Position: for HS5, the code is completed by a blank ON THE LEFT
8	1000	Alphanumerical	<b>Label</b> The label is completed by blanks ON THE RIGHT

The TXT file contents only one language. Language detection is done via analysis of the file name.

The file ***wcoomd-sh-nom-v4-2007-e3-200812-txt-en\_fr.zip*** contains 2 files:

***wcoomd-sh-nom-v4-2007-e3-200812-txt-en.txt***: the English version

***wcoomd-sh-nom-v4-2007-e3-200812-txt-fr.txt***: the French version

## Further Information and Specific Cases

Some labels are deliberately left empty in all languages: these are positions that are no longer used, or that will be subsequently used by the WCO.

The heading for all sub-positions (HS6) ending in "0" is prefixed by a single dash ("-"), with the exception of those ending in "00" which do not contain dashes.

By convention, sub-position labels ending in "00" end in a dot "."

The heading for all grouping positions (HS5) is prefixed by a single dash ("-").

By convention, grouping sub-position labels end in two dots ".."

The headings for all sub-positions (HS6) ending in a figure between "1" and "9" are prefixed by a double dash ("--").

### Examples (CSV version):

HS6 with a single dash:

In English:

4 2012.0204 2012.020410 - Carcasses and half-carcasses of lamb, fresh or chilled

In French:

4 2012.0204 2012.020410 - Carcasses et demi-carcasses d'agneau, fraîches ou réfrigérées

HS6 with no dash:

In English:

4 2012.0409 2012.040900 Natural honey.

In French:

4 2012.0409 2012.040900 Miel naturel.

HS5:

In English:

4 2012.0705 2012.07051 - Lettuce:

In French:

4 2012.0705 2012.07051 - Laitues :

HS6 with double dash:

In English:

4 2012.70193 2012.701932 -- Thin sheets

In French:

4 2012.70193 2012.701932 -- Voiles

### III- SH ALPHABETICAL INDEX FILES

#### XML files

##### Naming conventions:

Updated available versions:

**wccoomd-index-alphabetique-en.zip**

**wccoomd-index-alphabetique-fr.zip**

Every archive is composed of various XML files – one for every letter of the alphabet. The name of the file includes the prefix LETTRE (French) / LETTER (English) and the relevant letter of the alphabet.

##### File description

This file complies with the XSD diagram available on the WCO publications website.

URI: <http://wccoomdpublications.org/schemas/ia.xsd>

The file is encoded in UTF-8. It contains the node **wco-omd:letter**, which includes the attribute “id” – the letter from the alphabetical index. In every XML file, there is a single element of this type. Every node of type **wco-omd:letter** contains one or more **wco-omd:entry** elements, corresponding to entries from the alphabetical index.

Every **wco-omd:entry** element contains a **wco-omd:description** element. Every **wco-omd:description** element contains explanations for the entry in the index. It may contain one or more **wco-omd:word** elements (words from the alphabetical index) and sometimes **wco-omd:link** entries which represent references for other keywords inside the alphabetical index. An element of type **wco-omd:description** may also include a **wco-omd:greek** element, representing italic letters for chemical formula.

Further explanations for the **wco-omd:entry** are shown in the table below:

Attribute	Description	Possible values
idNomenclature	Reference to nomenclature	53.05 2906.21/2906.29 2906.11-2906.19
idNotesExplicatives	Reference to explanations	38.23 (29.05)   38.23 29 Gen (29.12)   (29.13)   29.05   29.06+29.07-29.11

Further explanations for the **wco-omd:description** are shown in the table below:

Element	Description	Possible values
wco-omd:description	Explanations together with <b>wco-omd:word</b> elements	Manila hemp, raw or processed but not spun, plus tow or waste
wco-omd:word	Contains keywords	ABALONE
wco-omd:link	Links to other keywords	MOLLUSC(s)
wco-omd:greek	Letters in italic for chemical product formula or Latin names	<i>p</i> -ACETAMIDOPHENOL <i>Musa textilis</i> Nee

## I. HS EXPLANATORY NOTES FILES

### XML and XHTML files

#### Naming conventions

The files are developed in line with the decisions of the Harmonised System Committee, which meets every 6 months and decides whether or not to amend the Explanatory Notes. The naming convention therefore takes into account the corrections made.

Available updated versions:

**XML update files:** contain only elements modified at the last committee meeting

**wcoomd-sh-ne-v5-2012-as-6-en-xml-201412.zip:** English version

**wcoomd-sh-ne-v5-2012-as-6-fr-xml-201412.zip:** French version

**wcoomd-sh-ne-v5-2012-as-5-es-xml-201406.zip:** Spanish version\*

**Complete XML files:** contain the latest version of all elements

**wcoomd-sh-ne-v5-2012-full-as-6-en-xml-201412.zip:** English version

**wcoomd-sh-ne-v5-2012-full-as-6-fr-xml-201412.zip:** French version

**wcoomd-sh-ne-v5-2012-full-as-5-es-xml-201406.zip:** Spanish version\*

\*The publication of the Spanish "n-1" version is synchronised with the publication of the official French and English "n" versions.

**Complete XMHTML files:** contain the latest version of all elements

**wcoomd-sh-ne-v5-2012-full-as-6-en-xhtml-201412.zip:** English version

**wcoomd-sh-ne-v5-2012-full-as-6-fr-xhtml-201412.zip:** French version

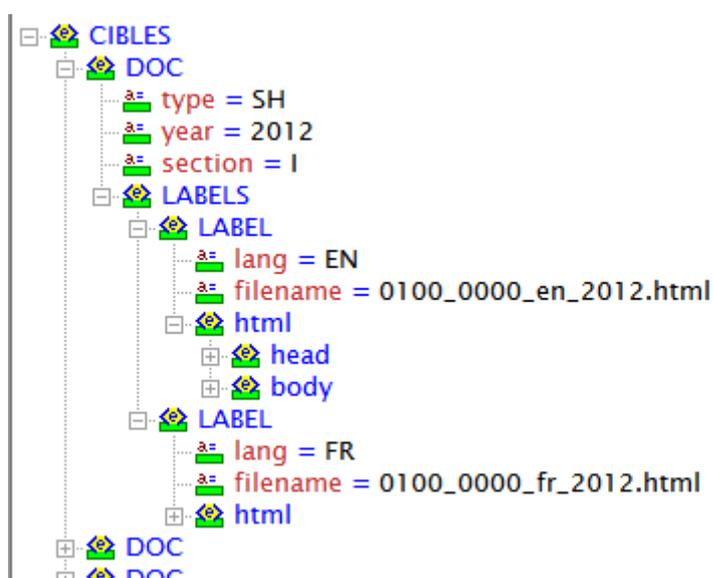
#### XML file description

This file complies with the XSD diagram available on the WCO website:

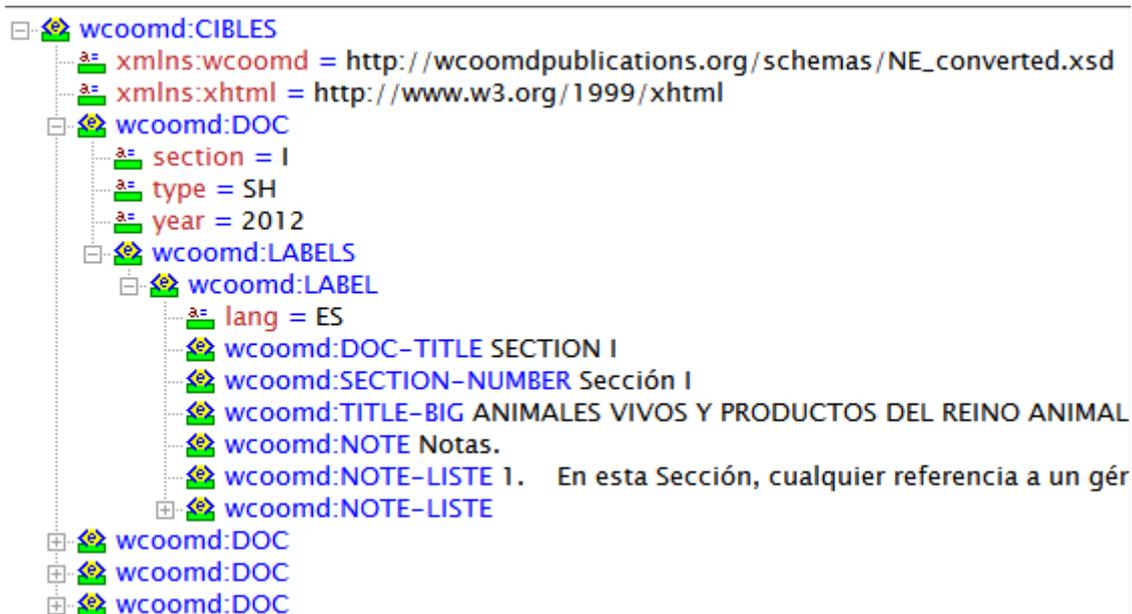
URL: [http://wcoomdpublications.org/schemas/NE\\_converted.xsd](http://wcoomdpublications.org/schemas/NE_converted.xsd)

The XML and XHTML versions are generally organised in the same way:

- A zip file contains the complete XML file and potentially a zip file of the images
- The root element is called "<TARGETS>"
- The classification is described by a "<DOC>" element (as many occurrences of <DOC> as there are Explanatory Notes)
- These languages are grouped in a "<LABELS>" element
- Strictly speaking, the Explanatory Note is contained in a "<LABEL>" element whose "Lang" attribute contains the language (EN = English, ES = Spanish, FR = French, IT = Italian, RU = Russian)
- In the case of XHTML, an "<html>" tag is used.



Tree of the XHTML version of the Explanatory Notes



Tree of the **XML** version of the Explanatory Notes

Further information on the attributes of the <DOC> element (relevant for **XML** and **XHTML** versions):

Attribute	Description	Possible values
Type	Type of content	Always equals "HS"
Year	HS version	2002, 2007, 2012, 2017
Section	The section number in Roman numerals	"I" to "XXI"
ID	The number of the chapter (2 digits) or of positions (4 digits). This attribute is absent for the section Explanatory Notes	"01" to "97" for the chapters "0101" to "9706" for the positions

In the case of **XHTML**, a "filename" attribute consisting of the original XHTML filename is entered at the "<LABEL>" tag level.

The **XHTML** version contains the formatting information. This is to be prioritised when the aim is to display Explanatory Notes on a webpage, for example. However, strictly speaking, it contains no indication of the structure of the Explanatory Note (titles, captions, paragraphs and lists are at the same level).

The **XML** version consists of end structured elements which can be used, for example, for a detailed indexation of contents (levels of titles, captions, notes, paragraphs, lists, etc.). The various levels available are described in the corresponding XSD diagram.

## II. HS CLASSIFICATION OPINIONS FILES

### XML files

#### Naming conventions

The files are developed in line with the decisions of the Harmonised System Committee, which meets every 6 months and decides whether or not to publish new Classification Opinions. The naming convention therefore takes into account the corrections made.

Available updated versions:

**Complete XML files:** contain the latest version of all elements

**wcoomd-sh-ac-v5-2012-full-as-6-en-xml-201412.zip:** English version

**wcoomd-sh-ne-v5-2012-full-as-6-fr-xml-201412.zip:** French version

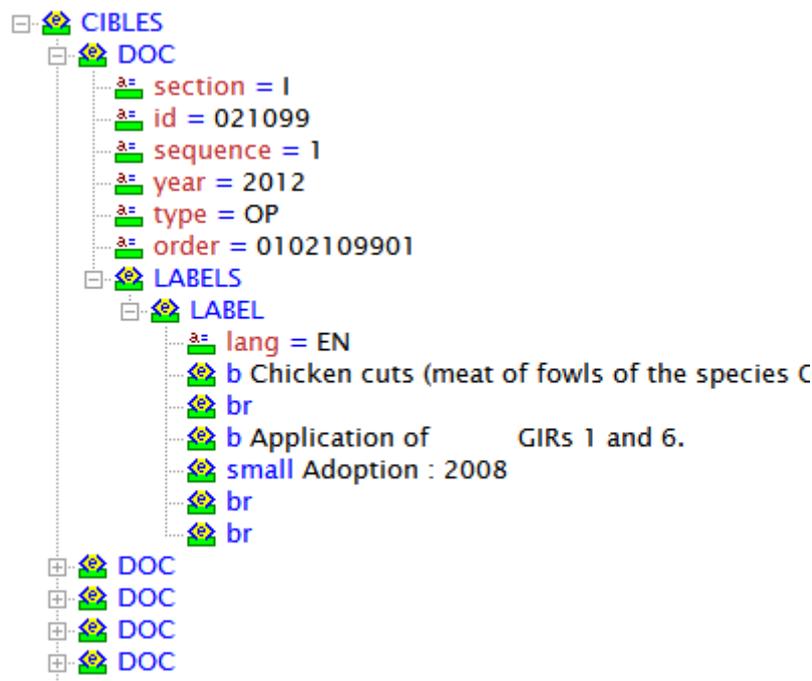
#### XML file description

This file complies with the XSD diagram available on the WCO website:

URL: <http://wcoomdpublications.org/schemas/AC.xsd>

The files are organised in the same way as the XML and XHTML files of the Explanatory Notes:

- A zip file contains the complete XML file and a zip file of the images
- The root element is called "<TARGETS>"
- Every Classification Option is described by a "<DOC>" element (as many occurrences of <DOC> as there are Explanatory Notes)
- These languages are grouped in a "<LABELS>" element
- Strictly speaking, the Classification Opinion is contained in a "<LABEL>" element whose "Lang" attribute contains the language (EN = English, ES = Spanish, FR = French, IT = Italian, RU = Russian)
- The Classification Opinion is in XHTML format (use of classic HTML formatting tags such as <b>, <i>, <br/>, etc.)



Tree of the **XML** version of the Classification Opinions

Further information on the attributes of the <DOC> element:

Attribute	Description	Possible values
Section	The section number in Roman numerals	"I" to "XXI"
ID	6-digit number of the sub-position (HS6) in question	For the first 9 chapters of the HS, the ID is prefixed by "0" Example: 021099
Sequence	The order number of the Classification Opinion in the sub-position	"1" to "99"
Year	HS version	2002, 2007, 2012, 2017
Type	Type of content	Always equals "OP"
Order	Complete number order of the Classification Opinion in the publication	It concerns the concatenation of the section (in 2-digit numerals), the sub-position ID (in 6-digit numerals) and the order number (in 2-digit numerals). Example for <b>0102109901</b> : Section = I ( <b>01</b> ) ID = <b>021099</b> Order = <b>01</b>

#### Further information on **image management**

A Classification Opinion may consist of 0 to n images. All the images are provided in a zip file accompanying the XML file.

To determine whether one or several images accompany a Classification Opinion, a search needs to be run in the zip file for files respecting the naming convention.

#### **Naming convention of the images:**

- The name of the images is composed of a string of 11 alphanumeric characters, a full stop and an extension.
- The 11-digit character string is made up of 3 parts:
  - o The first 6 digits correspond to the sub-position (HS 6), i.e. the ID attribute
  - o The subsequent 4 digits correspond to the order number (sequence attribute) potentially filled in as "0"
  - o The final digit corresponds to the order number of the image for this opinion, in alphabetical form and capitals (A=1, B=2, [...] Z=26)
- The extension is always in lower-case letters. Only 3 extensions are used: gif, jpg and png.

Example for the image **1902200001B.jpg**:

- The sub-position is "1902.20" (ID = 190220)
- It concerns the first Classification Opinion of this sub-position (sequence = 1)
- The image row is 2 (B = 2)

It therefore concerns the second illustration of the first Classification Opinion of the sub-position 1902.20.

# FICHIERS STRUCTURÉS DU SH

## I. CONVENTIONS DE NOMMAGE GÉNÉRALES DES FICHIERS STRUCTURÉS

L'OMD utilise une convention de nommage pour la diffusion des fichiers de ces publications. Le nom des fichiers (hors extensions) respecte cette convention.

Le nom des fichiers est composé de plusieurs caractéristiques séparées par un tiret (« - »). La taille du nom est variable en fonction des caractéristiques présentes, propres au fichier.

Les caractéristiques sont les suivantes (les caractéristiques obligatoires sont en gras) :

Caractéristique	Description	Valeurs possibles
<b>Nom de l'éditeur</b>	Comprend le nom de l'éditeur de la publication dont est issue le fichier	Toujours égal à « wcoomd »
<b>Collection</b>	Comprend l'abréviation de la collection dont dépend la publication dont est issu le fichier	SH : Système Harmonisé
<b>Publication</b>	Comprend l'abréviation de la publication dont est issu le fichier	NOM : Nomenclature NE : Notes Explicatives AC : Recueil d'avis de classement COM : Base de données des marchandises IA : Index alphabétique
<b>Version principale</b>	Comprend le préfixe « V » et la révision principale de la publication. Par exemple, pour le système harmonisé, la version communément appelée 2007 est la 4eme révision	V4 : Système Harmonisé 2007 V5 : Système Harmonisé 2012 V6 : Système Harmonisé 2017
<b>Date de version</b>	Il s'agit de la version « année » de la publication au format « AAAA ».	2007, 2012, 2017
Version subalterne	Comprend le préfixe « e » (pour erratum) et le numéro de l'erratum. Uniquement dans le cas de la nomenclature du système harmonisé.	Toutes les valeurs sont théoriquement possibles, le cas particulier « e0 » correspond à la version originale de la version principale
Numéro de révision	Comprend le préfixe « as » (pour « amending supplement ») puis le numéro du supplément. Uniquement dans le cas des Notes Explicatives et des Avis de Classement du système harmonisé ou le terme « full » s'il s'agit d'une version complète.	Toutes les valeurs sont théoriquement possibles. L'absence de cette caractéristique correspond à la version originale de la version principale.
<b>Date d'adoption</b>	Comprend la date, sous la forme « AAAAMM » (année sur 4 chiffres, mois sur 2 chiffres) d'adoption par les membres de l'OMD de la publication	Toutes les valeurs sont théoriquement possibles. Dans la pratique AAAA commence à partir de 2007 et MM est bien sûr compris entre « 01 » pour janvier et « 12 » pour décembre
<b>Format</b>	Comprend le format du fichier	CSV : Texte délimité TXT : Texte taille fixe DOC : Fichiers Word PDF : Fichiers PDF XML : Fichiers XML XHTML : Fichiers XHTML
<b>Langue(s)</b>	Comporte les langues présentes dans le fichier (chaque langue est séparée par le caractère « _ »).	En : Anglais Es : Espagnol Fr : Français It : Italien Ru : Russe
Simulation	Le suffixe « simulation » peut être utilisé pour qualifier des fichiers des tests.	

## Exemples concrets de nom de fichier conformes :

### ***wcoomd-sh-nom-v5-2012-e0-201006-csv-en\_fr.zip***

Fichier ZIP contenant la version originale de la nomenclature 2012 de juin 2010 au format CSV en français et en anglais.

### ***wcoomd-sh-nom-v4-2007-e3-200812-xml-en\_es\_fr\_it\_ru.zip***

Fichier ZIP contenant le 3eme erratum de la nomenclature 2007 de décembre 2008 au format XML en anglais, espagnol, français, italien et russe.

### ***wcoomd-sh-ne-v5-2012-full-201006-pdf\_fr.zip***

Fichier ZIP contenant la totalité des notes explicatives de la version 2012 du système harmonisé en français à juin 2010 au format PDF.

### ***wcoomd-sh-ne-v5-2012-as1-201112-doc\_en-simulation.zip***

Fichier ZIP contenant uniquement les notes explicatives corrigées par l'AS1 de la version 2012 du système harmonisé en anglais à décembre 2011 au format WORD, à des fins de simulations (données non officielles).

## II. FICHIERS DE LA NOMENCLATURE DU SH

### Fichiers ASCII délimité (CSV)

#### Conventions de nommage :

Versions à jour disponibles :

**wcoomd-sh-nom-v4-2007-e3-200812-csv-en\_fr.zip**

**wcoomd-sh-nom-v5-2012-e0-201006-csv-en\_fr.zip**

**wcoomd-sh-nom-v6-2017-e0-201606-csv-en\_fr.zip**

#### Description du fichier

Le caractère de tabulation ([TAB] CHR(9)) est utilisé comme délimiteur.

La première ligne contient les intitulés des colonnes.

Le fichier est encodé en UTF-8.

Position	Nom	Type	Description
1	Level	Numérique	Contient le niveau de l'entrée dans la classification 1 : Section 2 : Chapitre (HS2) 3 : Position (HS4) 4 : Sous-Position (HS6) ou Sous-Position de regroupement (HS5)
2	Father	Alphanumérique	Contient le code de la ligne de rattachement sous la forme : - <b>9999</b> (préfixe quatre chiffres arabes) : uniquement pour les sections, précise à quelle version de la nomenclature la section est rattachée (par exemple « 2012 ») - <b>9999.X</b> (suffixe sous la forme d'un chiffre romain) : uniquement pour les chapitres, précise à quelle section le chapitre est rattaché - <b>9999.99</b> (suffixe sous la forme de deux chiffres arabes) : uniquement pour les positions, précise à quel chapitre la position est rattachée - <b>9999.9999</b> (suffixe sous la forme de quatre chiffres arabes) : pour les sous-positions à 6 chiffres se terminant par « 0 » ou pour les sous-positions de regroupement. Précise la position de regroupement à laquelle la sous-position est rattachée. - <b>9999.99999</b> (suffixe sous la forme cinq chiffres arabes) : uniquement pour les sous-positions à 6 chiffres se terminant par un chiffre de « 1 » à « 9 ». Précise la sous-position de regroupement à laquelle la sous-position est rattachée
3	Hs_code	Alphanumérique	Contient le code SH de la ligne sous la forme : - <b>9999.X</b> (suffixe sous la forme d'un chiffre romain) : uniquement pour les sections - <b>9999.99</b> (suffixe sous la forme de deux chiffres arabes) : uniquement pour chapitres (HS2) - <b>9999.9999</b> (suffixe sous la forme de quatre chiffres arabes) : pour les positions (HS4). - <b>9999.99999</b> (suffixe sous la forme cinq chiffres arabes) : pour les sous-positions de rattachement (HS5) - <b>9999.999999</b> (suffixe sous la forme de six chiffres arabes) : pour les sous-positions (HS6)
4	LANG1	Alphanumérique	Intitulé de la ligne, dans la langue précisée par la première ligne
5	LANG2	Alphanumérique	Intitulé de la ligne, dans la langue précisée par la première ligne
...	...	...	...

Le nombre de colonnes varie en fonction du nombre de langues présentes dans le fichier. Le nombre de ligne est constant pour une version donnée de la nomenclature, quel que soit le nombre de langues présentes (à titre d'exemple, le fichier de la nomenclature 2012 originale comporte 7504 lignes).

La fin de chaque ligne est matérialisée par un retour chariot.

## Fichiers XML

### Conventions de nommage

Versions à jour disponibles :

**wcoomd-sh-nom-v4-2007-e3-200812-xml-en\_fr.zip**

**wcoomd-sh-nom-v5-2012-e0-201006-xml-en\_fr.zip**

**wcoomd-sh-nom-v6-2017-e0-201606-xml-en\_fr.zip**

### Description du fichier

Le fichier est conforme au schéma XSD disponible sur le site web des publications de l'OMD.

URI : [http://wcoomdpublications.org/schemas/HS\\_converted.xsd](http://wcoomdpublications.org/schemas/HS_converted.xsd)

Le fichier est encodé en UTF-8. Il contient autant de noeuds « NOMENCLATURE » qu'il y a de lignes et, pour chacune, autant de nœud « LABEL » qu'il y a de langues.

Chaque nœud « LABEL » reçoit l'intitulé de la ligne dans la langue indiquée dans l'attribut « lang ».

Les langues sont conformes à la codification ISO 639-1 (alpha-2) :

en : Anglais  
 es : Espagnol  
 fr : Français  
 it : italien  
 ru : Russe

### Précisions sur les attributs du nœud « NOMENCLATURE »

Le nœud <NOMENCLATURE> comporte plusieurs attributs dont certains sont optionnels (voir le schéma XSD). Le tableau ci-dessous précise certaines caractéristiques pour faciliter la compréhension du fichier XML.

Les attributs présentés **en gras** sont obligatoires.

Attribut	Description	Valeurs possibles
<b>Year</b>	Contient la version principale du SH (année) à laquelle se réfère la ligne. L'année est encodé en numérique sur 4 caractères (9999).	2007 2012 2017
<b>Section</b>	Contient le numéro de la section, en chiffres romains, à laquelle est rattachée la ligne.	« I » à « XXI »
<b>Order</b>	Contient le numéro d'ordre logique dans la classification qui ne contient que des chiffres numériques. Chaque niveau est représenté par 2 digits, sauf dans le cas des sous-positions de regroupement. Les sections ont 2 digits. Les chapitres ont 4 digits Les positions ont 6 digits Les sous-positions ont 8 digits Les sous-positions de regroupement ont 7 digits.	<b>Chiffres 1 et 2</b> : le numéro de section en chiffres arabes, de 01 à 21 <b>Chiffres 3 et 4</b> : le numéro de chapitre de 01 à 97 <b>Chiffres 5 et 6</b> : le numéro de position (variable) <b>Chiffres 7 et 8</b> : le numéro de sous-position (variable) Lorsque <b>le chiffre 8 seul est absent</b> , il s'agit d'une sous-position de regroupement
<b>Level</b>	Contient le niveau de l'entrée dans la classification 1 : Section 2 : Chapitre (HS2) 3 : Position (HS4) 4 : Sous-Position (HS6) ou Sous-Position de regroupement (HS5)	
<b>Chapter</b>	Contient le numéro du chapitre sur 2 digits (équivalent aux chiffres 3 et 4 de l'attribut «	<b>01 à 97</b>

	order »)	
Heading	Contient le numéro de position dans le chapitre sur 2 digits (équivalent aux chiffres 5 et 6 de l'attribut « order »)	<b>Variable</b>
Subheading	Contient le numéro de sous-position dans la position	<b>Variable</b>

## **Fichiers ASCII à taille fixe (TXT)**

### **Conventions de nommage**

Versions à jour disponibles :

***wcoomd-sh-nom-v4-2007-e3-200812-txt-en\_fr.zip***

***wcoomd-sh-nom-v5-2012-e0-201006-txt-en\_fr.zip***

***wcoomd-sh-nom-v6-2017-e0-201606-txt-en\_fr.zip***

### **Description du fichier**

Chaque ligne a une taille fixe de 1008 caractères. La ligne est terminée par un retour chariot.

Le fichier est encodé en UTF-8.

La première ligne contient la date d'adoption de la version fournie.

<b>Position</b>	<b>Longueur</b>	<b>Type</b>	<b>Description</b>
1	1	Numérique	Contient le niveau de l'entrée dans la classification 1 : Section 2 : Chapitre (HS2) 3 : Position (HS4) 4 : Sous-Position (HS6) ou Sous-Position de regroupement (HS5)
2	6	Alphanumérique	<b>Code SH</b> <b>Section</b> : le code est fourni en numérique (à traduire en chiffres romains par vos soins) sur deux digits, complété par des blancs A GAUCHE. <b>Chapitre</b> : le code est fourni sur deux digits, complété par des blancs A GAUCHE <b>Position</b> : le code HS4 est fourni sur quatre digits, complétés par des blancs A GAUCHE Sous-Position / Sous-Position de regroupement : pour les HS5, le code est complété par un blanc A GAUCHE
8	1000	Alphanumérique	<b>Libellé</b> Le libellé est complété par des blanc A DROITE

Le fichier TXT ne comporte qu'une seule langue. La discrimination de la langue se fait par analyse du nom du fichier. Le fichier ***wcoomd-sh-nom-v4-2007-e3-200812-txt-en\_fr.zip*** contient 2 fichiers :

***wcoomd-sh-nom-v4-2007-e3-200812-txt-en.txt*** : la version anglaise

***wcoomd-sh-nom-v4-2007-e3-200812-txt-fr.txt*** : la version française

## Précisions et Cas particuliers

Certains libellés sont volontairement vides dans toutes les langues : il s'agit de positions qui ne sont plus utilisées ou qui seront utilisées ultérieurement par l'OMD.

L'intitulé de toutes les sous-positions (HS6) qui se terminent par « 0 » est préfixé par un tiret simple (« - ») à l'exception de celles qui se terminent par « 00 » pour lesquelles il n'y a pas de tiret.

Par convention, les libellés des sous-positions qui se terminent par « 00 » sont terminés par un point « . »

L'intitulé de toutes les positions de regroupement (HS5) est préfixé par un tiret simple (« - »).

Par convention, les libellés des sous-positions de regroupement sont terminée par deux points « : ».

L'intitulé de toutes les sous-positions (HS6) qui se terminent par un chiffre compris entre « 1 » et « 9 » est préfixé par un double tiret (« -- »).

### **Exemples (version CSV) :**

#### **HS6 avec un seul tiret :**

En anglais :

4 2012.0204 2012.020410 - Carcasses and half-carcasses of lamb, fresh or chilled

En français :

4 2012.0204 2012.020410 - Carcasses et demi-carcasses d'agneau, fraîches ou réfrigérées

#### **HS6 sans tiret :**

En anglais :

4 2012.0409 2012.040900 Natural honey.

En français :

4 2012.0409 2012.040900 Miel naturel.

#### **HS5 :**

En anglais :

4 2012.0705 2012.07051 - Lettuce :

En français :

4 2012.0705 2012.07051 - Laitues :

#### **HS6 avec double tiret :**

En anglais :

4 2012.70193 2012.701932 -- Thin sheets (voiles)

En français :

4 2012.70193 2012.701932 - Voiles

### III. FICHIERS DE L'INDEX ALPHABÉTIQUE DU SH

#### Fichiers XML

##### Conventions de nommage

Versions à jour disponibles:

**wcoomd-index-alphabetique-en.zip**

**wcoomd-index-alphabetique-fr.zip**

Chaque archive contient des fichiers XML- une pour chaque lettre de l'alphabet. Le nom d'un fichier est composé d'un préfix LETTRE (français) / LETTER (anglais) et de la lettre de l'index alphabétique.

##### Description du fichier

Le fichier est conforme au schéma XSD disponible sur le site web de l'OMD:

URI: <http://wcoomdpublications.org/schemas/ia.xsd>

Le fichier est encodé en UTF-8. Il contient des nœuds **wco-omd:letter**, qui possèdent l'attribut **id** – la lettre de l'index alphabétique. Il y a un seul élément **wco-omd:letter** dans un fichier XML. Chaque nœud **wco-omd:letter** contient plusieurs éléments **wco-omd:entry** correspondant à une entrée de l'index alphabétique.

Chaque élément **wco-omd:entry** contient un élément **wco-omd:description**. Chaque élément **wco-omd:description** peut contenir à son tour un ou plusieurs éléments du type **wco-omd:word** et du texte explicatif pour l'entrée de l'index. Un élément **wco-omd:description** peut aussi contenir des références correspondantes à des liens vers d'autres mots clé **wco-omd:link**, mais aussi des éléments correspondants aux lettres en italique pour les formules chimiques : **wco-omd:greek**.

Précisions sur les attributs du nœud **wco-omd:entry** :

Attribut	Description	Valeurs possibles
idNomenclature	Référence de la nomenclature	53.05 2906.21/2906.29 2906.11-2906.19
idNotesExplicatives	Références des notes explicatives	38.23 (29.05)   38.23 29 Gen (29.12)   (29.13)   29.05   29.06+29.07-29.11

Précisions sur le nœud **wco-omd:description** :

Élément	Description	Valeurs possibles
wco-omd:description	Contenu explicatif avec wcoomd:word	Utilisés principalement pour l'alimentation humaine
wco-omd:word	Contient des mots clé	comestibles
wco-omd:link	Liens vers d'autres mots clé	VIANDES
wco-omd:greek	Correspond à des lettres en italique pour les formules chimiques ou des noms latins	<i>n</i> -propyle <i>Musa textilis</i> Nee

## IV. FICHIERS DES NOTES EXPLICATIVES DU SH

### Fichiers XML et XHTML

#### Conventions de nommage

Les fichiers évoluent au rythme de décisions du Comité du Système Harmonisé qui se réunit tous les 6 mois et décide ou non de procéder à des amendements des Notes Explicatives. Le convention de nommage tient donc compte des corrections réalisées.

Versions à jour disponibles:

**Fichiers XML de mise à jour** : ne contiennent que les éléments modifiés lors du dernier comité

**wcoomd-sh-ne-v5-2012-as-6-en-xml-201412.zip** : version anglaise

**wcoomd-sh-ne-v5-2012-as-6-fr-xml-201412.zip** : version française

**wcoomd-sh-ne-v5-2012-as-5-es-xml-201406.zip** : version espagnole \*

**Fichier XML complets** : contiennent la dernière version de l'intégralité des éléments

**wcoomd-sh-ne-v5-2012-full-as-6-en-xml-201412.zip** : version anglaise

**wcoomd-sh-ne-v5-2012-full-as-6-fr-xml-201412.zip** : version française

**wcoomd-sh-ne-v5-2012-full-as-5-es-xml-201406.zip** : version espagnole \*

\* La publication de la version « n-1 » espagnole est synchronisée avec la publication des versions officielles « n » française et anglaise.

**Fichier XMHTML complets** : contiennent la dernière version de l'intégralité des éléments

**wcoomd-sh-ne-v5-2012-full-as-6-en-xhtml-201412.zip** : version anglaise

**wcoomd-sh-ne-v5-2012-full-as-6-fr-xhtml-201412.zip** : version française

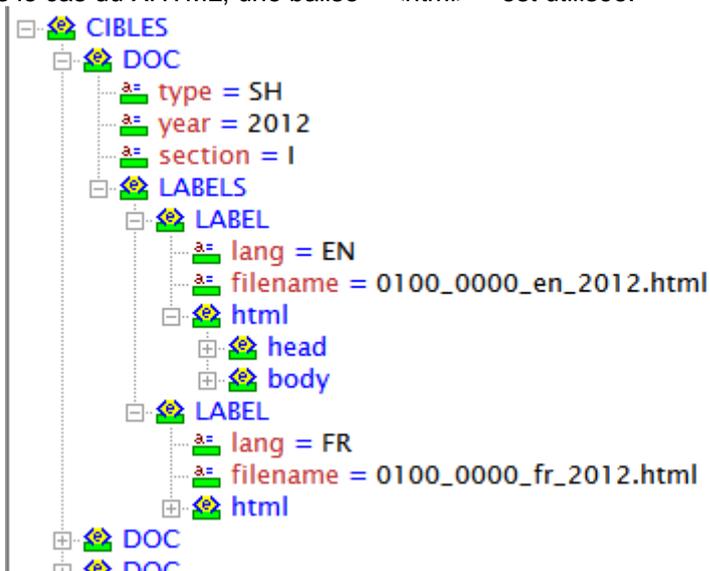
#### Description du fichier XML

Le fichier est conforme au schéma XSD disponible sur le site web de l'OMD:

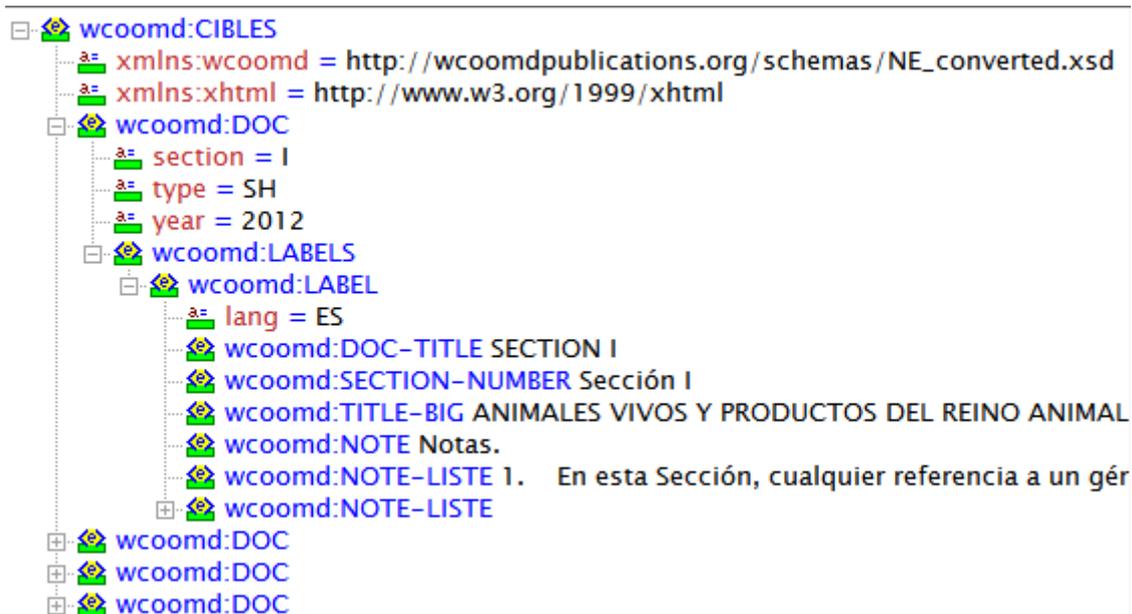
URI: [http://wcoomdpublications.org/schemas/NE\\_converted.xsd](http://wcoomdpublications.org/schemas/NE_converted.xsd)

Les versions XML et XHTML sont globalement organisées de la même manière :

- Un fichier ZIP contient le fichier XML complet et un fichier ZIP éventuel comportant les images
- L'élément racine s'appelle « <CIBLES> »
- La classification est décrite par un élément « <DOC> » (autant d'occurrences <DOC> qu'il y a de Notes Explicatives)
- Les langues sont éventuellement regroupées dans un élément « <LABELS> »
- La Note Explicative proprement dite est contenu dans un élément « <LABEL> » dont l'attribut « Lang » contient la langue (EN = anglais, ES = espagnol, FR = français, IT = italien, RU = russe)
- Dans le cas du XHTML, une balise « <html> » est utilisée.



Arbre de la version XHTML des Notes explicatives



Arbre de la version **XML** des Notes Explicatives

Précisions sur les attributs de l'élément <DOC> (valable pour les versions **XML** et **XHTML**) :

Attribut	Description	Valeurs possibles
type	Type de contenu	Vaut toujours « SH »
year	Version du SH	2002, 2007, 2012, 2017
section	Le numéro de section en chiffres romains	« I » à « XXI »
id	Le numéro de chapitre (sur 2 digits) ou de positions (sur 4 digits). Cet attribut est absent pour les Notes Explicatives de section	« 01 » à « 97 » pour les chapitres « 0101 » à « 9706 » pour les positions

Dans le cas du **XHTML**, un attribut « filename » comportant le nom du fichier XHTML original est renseigné au niveau de la balise « <LABEL> ».

La version **XHTML** comporte les informations de mise en forme. Elle est à privilégier lorsque l'objectif est de réaliser un affichage des Notes Explicatives dans une page Web par exemple. Elle ne comporte toutefois aucune indication de structuration de la Note Explicative proprement dite (titres, sous-titres, paragraphes et listes sont au même niveau).

La version **XML** comporte des éléments de structuration fin qui peuvent être utilisés, par exemple, pour une indexation fine des contenus (niveaux de titres, sous-titres, notes, alinéas, listes...). Les différents niveaux disponibles sont décrits dans le schéma XSD correspondant.

## V. FICHIERS DES AVIS DE CLASSEMENT DU SH

### Fichiers XML

#### Conventions de nommage

Les fichiers évoluent au rythme de décisions du Comité du Système Harmonisé qui se réunit tous les 6 mois et décide ou non de procéder à la publication de nouveaux Avis de Classement. Le convention de nommage tient donc compte des corrections réalisées.

Versions à jour disponibles:

**Fichier XML complets** : contiennent la dernière version de l'intégralité des éléments

**wcoomd-sh-ac-v5-2012-full-as-6-en-xml-201412.zip** : version anglaise

**wcoomd-sh-ne-v5-2012-full-as-6-fr-xml-201412.zip** : version française

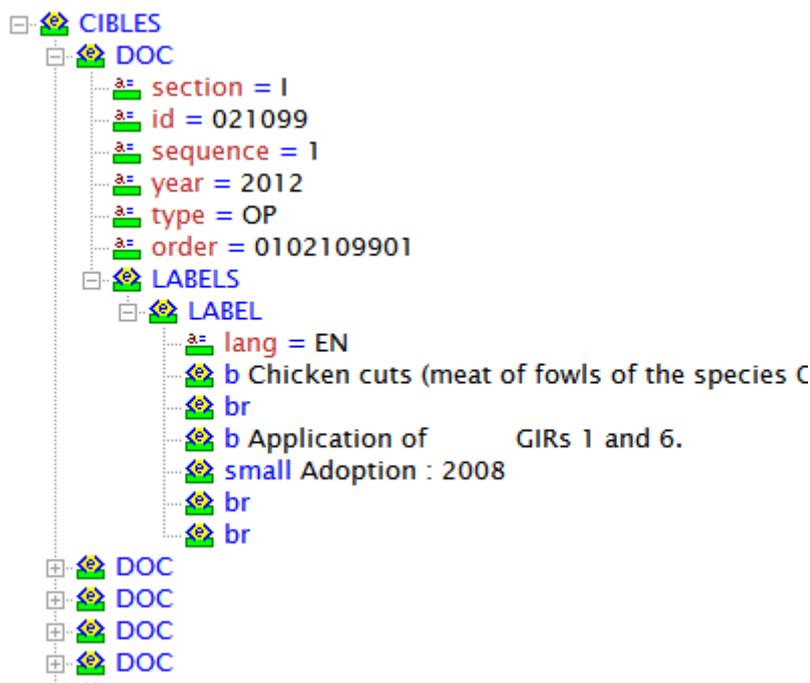
#### Description du fichier XML

Le fichier est conforme au schéma XSD disponible sur le site web de l'OMD:

URI: <http://wcoomdpublications.org/schemas/AC.xsd>

Les fichiers sont organisés de la même manière que les fichiers XML et XHTML des Notes Explicatives :

- Un fichier ZIP contient le fichier XML complet et un fichier ZIP comportant les images
- L'élément racine s'appelle « <CIBLES> »
- Chaque Avis de Classement est décrit par un élément « <DOC> » (autant d'occurrences <DOC> qu'il y a de Notes Explicatives)
- Les langues sont éventuellement regroupées dans un élément « <LABELS> »
- L'Avis de Classement proprement dit est contenu dans un élément « <LABEL> » dont l'attribut « Lang » contient la langue (EN = anglais, ES = espagnol, FR = français, IT = italien, RU = russe)
- L'Avis de Classement est au format XHTML (utilisation de balises HTML classiques de mise en forme comme <b>, <i>, <br/>...)



Arbre de la version XML des Avis de Classement

Précisions sur les attributs de l'élément <DOC> :

Attribut	Description	Valeurs possibles
section	Le numéro de section en chiffres romains	« I » à « XXI »
id	Numéro de sous-position concernée sur 6 digits (HS6)	Pour les 9 premières chapitres du SH, l'id est préfixé par « 0 » Exemple : 021099
sequence	Le numéro d'ordre de l'Avis de Classement dans la sous-position	« 1 » à « 99 »
year	Version du SH	2002, 2007, 2012, 2017
type	Type de contenu	Vaut toujours « OP »
order	Numéro d'ordre complet de l'Avis de Classement dans la publication	Il s'agit de la concaténation de la section (en numérique sur 2 digit), de l'id de sous-position (en numérique sur 6 digits) et du numéro d'ordre (en numérique sur 2 digits). Exemple pour <b>0102109901</b> : section = I ( <b>01</b> ) id = <b>021099</b> order = <b>01</b>

### Précisions sur la gestion des images

Un Avis de Classement peut comporter de 0 à n images. Toutes les images sont fournies dans un fichier ZIP accompagnant le fichier XML.

Pour déterminer si une ou plusieurs images accompagnent un Avis de Classement, il faut rechercher dans ce fichier ZIP la présence de fichier(s) respectant la convention de nommage.

#### Convention de nommage des images :

- Le nom des images est composé d'une chaîne de caractères de 11 digits alphanumériques, d'un point et d'une extension.
- La chaîne de caractères de 11 digits est construite de 3 parties :
  - o Les 6 premiers digits correspondent à la sous-position (HS 6) c'est-à-dire à l'attribut id
  - o Les 4 digits suivant correspondent au numéro d'ordre (attribut sequence) complétés éventuellement par « 0 »
  - o Le dernier digit correspond au numéro d'ordre de l'image pour cet avis, sous forme alphabétique en majuscule (A=1, B=2, ..., Z=26)
- L'extension est toujours en minuscule. Seules 3 extensions sont utilisées : gif, jpg et png.

Exemple pour l'image **1902200001B.jpg** :

- La sous position est « 1902.20 » (id = 190220)
- Il s'agit du premier Avis de Classement de cette sous position (sequence = 1)
- Le rang de l'image est 2 (B = 2)

Il s'agit donc de la 2ème illustration du 1<sup>er</sup> Avis de Classement de la sous-position 1902.20.