

Yellow Book 2

U. S. Patent Images/TIFF

**United States Patents
(Grants and Published Applications)
Delivered as
CCITT Group 4 Facsimile Images**

**February 1, 2005
Revised August 27, 2012**

**United States Patent & Trademark Office
Electronic Information Products Division**

August 27, 2012

Update to U.S. Patent Image/TIFF (a.k.a. Yellow Book 2)

Page 9 : Update DTD comments as follows to identify kind codes for the America Invents Act (AIA) Certificates.

2012-08-27 Ed Johnson

Identify Kind Codes "C", "F", "J", "K" and "O" as representing the following Certificates"

Cn - Reexamination Certificate issued on or after *January 2, 2001*

NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

There are three (3) types of Reexamination Certificates for kind code "C":

- (1). Ex Parte Reexamination Certificate under 35 U.S.C. 307
- (2). Ex Parte Reexamination Certificate under 35 U.S.C. 257
- (3). Inter Partes Reexamination Certificate under 35 U.S.C. 316

Fn - Supplemental Examination Certificate published after September 16, 2012

NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

Jn - Post Grant Review Certificate published after September 16, 2012

NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

Kn - Inter Partes Review Certificate published after September 16, 2012

NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

On - Derivation Certificate published after March 16, 2013

~~~~~

August 27, 2012

## Update to U.S. Patent Image/TIFF (a.k.a. Yellow Book 2)

Page 17: Update Table 1a U.S. Patent Grant Patent Numbers to include Certificates

Page 19 and Page 20: Update Table 2 - U.S. Patent Grants and Patent Published Applications – Kind Codes as follows to identify kind codes for the America Invents Act (AIA) Certificates.

Cn - Reexamination Certificate issued on or after *January 2, 2001*

NOTE: “n” represents a value 1 through 9 denoting the publication level.

There are three (3) types of Reexamination Certificates for kind code “C”:

- (1). Ex Parte Reexamination Certificate under 35 U.S.C. 307
- (2). Ex Parte Reexamination Certificate under 35 U.S.C. 257
- (3). Inter Partes Reexamination Certificate under 35 U.S.C. 316

Fn – Supplemental Examination Certificate published after September 16, 2012

NOTE: “n” represents a value 1 through 9 denoting the publication level.

Jn - Post Grant Review Certificate published after September 16, 2012

NOTE: “n” represents a value 1 through 9 denoting the publication level.

Kn - Inter Partes Review Certificate published after September 16, 2012

NOTE: “n” represents a value 1 through 9 denoting the publication level.

On - Derivation Certificate published after March 16, 2013

NOTE: “n” represents a value 1 through 9 denoting the publication level.

~~~~~

February 1, 2012

Update to U.S. Patent Image/TIFF (a.k.a. Yellow Book 2)

Page 1:

2. Summary

Change the external media for dissemination of Patent Images/TIFF files from DLT cartridge to optical disc (Blu-ray or DVD).

3. Organization of Optical Disc Content

Change the title of item 3. from **Organization of DLT Contents to Organization of Optical Disc Content**

3.1 Media ID File

Change the title of item 3.1 from **Tape ID File to Media ID File**

Page 2:

Change the name of the Series Code table from **Tape Series Code to Media Series Code**

Identify all the types of optical disc products to include the separate optical disc for Reexamination Certificates that began with the issue date of October 4, 2011.

3.2 Content List File

Change the title of item 3.2 from **Tape Content List (TCL) File to Content List File (a.k.a. TCL)**

Page 3, 4:

Examples of an updated Content List (TCL) for:

- Patent Grants
- Certificates of Correction
- Reexamination Certificates
- Published Applications

February 1, 2012

Update to U.S. Patent Image/TIFF (a.k.a. Yellow Book 2) – (continued)

Page 4, 5, 6, 7:

Examples of updated:

Directory Structure Patent Grant Publications
Directory Structure Hierarchy Patent Grant Publications
Directory Structure Example - Patent Grant Publications

Directory Structure Patent Applications Publications
Directory Structure Hierarchy Patent Applications Publication
Directory Structure Example – Patent Published Applications

Page 12:

Include the following comments in the DTD:

NOTE: Fn present for a Reexamination Certificate of a Reissue published prior to January 12, 2010.

A1 - Utility Patent Grant issued prior to January 2, 2001

Page 13:

Update the following DTD comment to identify proper date range for “X” patents:

I1 - "X" Patents issued from July 31, 1790 to July 3, 1836.

I2 - "X" Reissue Patents issued from July 7, 1836 to June 13, 1848.

Include the following comments in the DTD:

X6 - Certificates of Correction.

Page 17:

Identify in Table 1a the content of the patent number field for all types of Certificates.

Identify in Table 1a that all types of Statutory Invention Registrations will contain “H” in position-1 of the patent number field.

February 1, 2012

Update to U.S. Patent Image/TIFF (a.k.a. Yellow Book 2) – (continued)

Page 17 (continued):

Identify proper date range for “X” patents

Include the appearance of Reissue “RX” patent numbers.

Page 18:

Include the appearance of fraction patent numbers.

Include Table 1b - U.S. Patent Application Publication Number – that identifies the appearance of U.S. Patent Application Publication Numbers

Page 19:

Include the appearance of the “Bn” kind code.

Include the appearance of the “Fn” kind code.

Include the date range for the “I2” kind code.

~~~~~

## **1. Background**

The original USPTO implementation of the World Intellectual Property Office (WIPO) Standard ST.33 is known as U.S. Patent Images/TIFF Yellow Book. ST.33 provides a proprietary header for CCITT Group 4 compressed raster images. This proprietary Yellow Book was discontinued the week ending June 18, 2004.

U.S. Patent Images/TIFF Yellow Book 2 (placed into production the week beginning June 21, 2004, uses a TIFF header for CCITT Group 4 compressed raster images of the pages in the patent document, accompanied by an XML instance with additional metadata for each patent document. Yellow Book 2 is based on WIPO Standards ST.33, ST.35, and current USPTO practice.

## **2. Summary**

U.S. Patent Images/TIFF (a.k.a. Yellow Book 2) consists of United States weekly Patent Grant Publications and weekly Patent Application Publications, as well as Certificates of Correction and Reexamination Certificates delivered as CCITT Group 4 facsimile images enclosed in TIFF headers. Each page of a patent document is in a single TIFF file. The files are organized into directories, one directory per patent document. The external media for dissemination of U.S. Patent Images/TIFF files (Patent Grant Publications and Patent Application Publications) will be optical disc (Blu-ray or DVD). Also included on each optical disc is a Media ID file and a Content List file that identifies all document numbers.

## **3. Organization of Optical Disc Content**

### **3.1. Media ID File**

The Media File ID file contains a 1-position Media Series Code (as defined in the following table) followed by a 5-position numeric serial number and the file extension .tid. Example of a Media ID File: Xnnnnn.tid, where "X" represents a Media Series Code identified in the following table. The Media ID file contains no data.

| <b>Media Series Code</b> |                                                                                                                                                     |
|--------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| A                        | Patent Grants prior to June 4, 2002                                                                                                                 |
| B                        | Replacement Documents captured from paper,                                                                                                          |
| B                        | Certificates of Correction captured from paper.                                                                                                     |
| G                        | Patent Grants                                                                                                                                       |
| M                        | Replacement Documents, Certificates of Correction, Reexamination<br>Certificates dissemination to the patent examining search system <sup>1</sup> . |
| P                        | Patent Application Publications                                                                                                                     |
| R                        | Reexamination Certificates – Disseminated on separate weekly optical<br>disc beginning October 4, 2011                                              |
| R5                       | Certificates of Extension                                                                                                                           |
| Z                        | Certificates of Correction captured electronically                                                                                                  |
| Z                        | Certificates of Correction (Patent Term Adjustment)<br>captured electronically                                                                      |

### 3.2. Content List File (a.k.a. TCL)

The file name for the Content List File (a.k.a.TCL) will be:

yyyymmdd.contents

For weekly publication of patent grants (G) and patent application publications (P), yyyy is the year, mm is the month and dd the day of the month, representing the issue/publication date of the patent documents on the file. For all other types of content, yyyymmdd is the date the file was created.

The Content List will identify each patent grant, patent application publication, certificate of correction and reexamination certificate present on the appropriate file. The Content List file will be in ASCII format, tab delimited. Each document in a Content List file will contain the document ID, the current kind code, the issue/publication date, and page count.. The data fields will be separated by a space - “hex 20” and each record/document terminated by a linefeed character - “hex 0A”,

---

<sup>1</sup> The Maintenance “M” file for the patent examining search system will contain:  
Replacement documents with the <status> field containing “RESCAN”.  
Certificates of Correction (The original document followed by the Certificate of Correction with the <status> field containing “COC”).  
Reexamination Certificates (The original document followed by the Reexamination Certificate with the <status> field containing “REEXAM”).



**Example of a Content List (TCL) for Patent Grants:**

**Grant Yellow Book 2**  
**TCL for Media ID Number: Gnnnnn**  
**Issue Date: mm/dd/yyyy**

| <b>Patent Number</b> | <b>Kind</b> | <b>Pages</b> |
|----------------------|-------------|--------------|
| D0653786             | S1          | 7            |
| H0002266             | H1          | 5            |
| PP022481             | P2          | 4            |
| RE043157             | E1          | 28           |
| 08108944             | B1          | 12           |

**Total Patents:** n,nnn  
**Total Pages:** nn,nnn  
**Total Bytes:** n,nnn,nnn,nnn

**Example of a Content List (TCL) for Certificates of Correction:**

**Certificate of Correction Yellow Book 2**  
**TCL for Media ID Number: Znnnnn**  
**Issue Date: mm/dd/yyyy**

| <b>Patent Number</b> | <b>Kind</b> | <b>Pages</b> |
|----------------------|-------------|--------------|
| RE042013             | X6          | 2            |
| 06667344             | X6          | 1            |

**Total Patents:** n,nnn  
**Total Pages:** nn,nnn  
**Total Bytes:** n,nnn,nnn,nnn

**Example of a Content List (TCL) for Reexamination Certificates:**

**Grant Yellow Book 2**  
**TCL for Media ID Number: Rnnnnn**  
**Issue Date: mm/dd/yyyy**

| <b>Patent Number</b> | <b>Kind</b> | <b>Pages</b> |
|----------------------|-------------|--------------|
| C7389718             | C1          | 2            |
| C7596474             | C1          | 9            |

**Total Patents:** n  
**Total Pages:** nn  
**Total Bytes:** n,nnn,nnn

## Example of a Content List for Patent Applications:

**Application Yellow Book 2**  
**TCL for Media ID number: Pnnnnn**  
**Projected Publication Date: mm/dd/yyyy**

| <b>Appl. Number</b> | <b>Kind</b>    | <b>Pages</b> |
|---------------------|----------------|--------------|
| 0036602             | A1             | 53           |
| 0036603             | A1             | 14           |
| 0036604             | P1             | 4            |
| <b>Total App's:</b> | n,nnn          |              |
| <b>Total Pages:</b> | nnn,nnn        |              |
| <b>Total Bytes:</b> | n, nnn,nnn,nnn |              |

### 3.3. Images and Metadata File

Documents are grouped under a directory that is at the root of the directory structure. The directory will be named as follows:

yyyy-ww

where yyyy is the year and ww is two digit week of the year that the documents wher created or modified.

## 4. Document Image Pages

### 4.1. Directory Structure Patent Grant Publications

A directory structure will be created for each Patent Grant Publication, to store the page images (TIFF files) and the document-level metadata (XML instance file)

#### 4.1.1 Directory Structure Hierarchy Patent Grant Publications

The hierarchy of the directory structure containing patent grants will be:

*Root\_Directory\_Name*

The *Root\_Directory\_Name* will contain YYYY-WW

where YYYY is the year and WW is two digit week of the year that the documents were created or modified.

Following the root directory will be 8-position patent numbers intended as follows to ensure that there are no more than 1,000 subdirectories in a directory:

1) – a two-position subdirectory identifying position-1 and position-2 of the patent number(s).

- Two-positions numeric with leading zero - Utility Patents
- “D0” - Design Patents
- “PP” – Plant Patents
- “RE” – Reissue Patents
- “H0” - Statutory Invention Registration (SIR)

2) – a three-position subdirectory identifying position-3, position-4 and position-5 of the patent number(s).

3) – a three-position subdirectory identifying position-6, position-7 and position-8 of the patent number(s).

#### 4.1.2 Directory Structure Example - Patent Grant Publications

A root directory listing for patent grants published the 2<sup>nd</sup> week of 2012,  
Issue date - 20120110  
G00001.tid  
20020319.contents  
2002-12

A directory listing for new patent grants published in the 12<sup>th</sup> week of 2002, showing the subdirectory for document 6,342,021 follows:

```
2002-02
| -06
| | --245
| | | --001
| | | --002
... ..
| | --342
| | | --021
| | | |--00000001.tif
| | | |--00000002.tif
... ..
| | | |--00003999.tif
| | | |--us-patent-image.xml
| | |--022
... ..
```

### **4.1.3 Image Page(s) .tif Files for Patent Grant Publications**

The TIFF file name for each image page will be:

nnnnnnnn.tif

where nnnnnnnn is an eight-character field containing the page number, right-aligned with leading zeros. The page number represents the sequence of the image page within the document.

## **4.2 Directory Structure Patent Applications Publications**

A directory structure will be created for each Patent Application Publication, to store the page images (TIFF files) and the document-level metadata (XML instance file).

### **4.2.1 Directory Structure Hierarchy Patent Applications Publication**

The hierarchy of the directory structure containing patent application publications will be:

*Root\_Directory\_Name*

The Root\_Directory\_Name will contain YYYY-WW

where YYYY is the year and WW is two digit week of the year that the documents were created or modified.

Following the root directory will be 15-position published application numbers and kind codes intended as follows to ensure that there are no more than 1,000 subdirectories in a directory:

- 1) – a two-position subdirectory (position-1, position-2) containing “US” identifying the United States as the publishing country.
- 2) – a four-position subdirectory (position-3, position-4, position-5, position-6) identifying the year (yyyy) of publication.
- 3) – a four-position subdirectory (position-7, position-8, position- 9, position-10) of the published application number.
- 3) – a three-position subdirectory (position-11, position-12, position- 13) of the published application number.
- 4) – a two-position subdirectory (position-14, position-15) containing the kind code of the published application.

## 4.2.2 Directory Structure Example – Patent Published Applications

A root directory listing for patent published applications published the 3rd week of 2002,  
Issue date - 20020117

```
P12345.tif  
20020117.contents  
2002-03
```

A directory listing for new patent published applications published in the 3<sup>rd</sup> week of 2002, showing  
the subdirectory for document US20020005880A1:

```
2002-03  
|-US  
|  |--2002  
|  |  |--0000  
|  |  |--0001  
  
... ..  
|  |  |--0005  
|  |  |  |--001  
|  |  |  |--002  
  
... ..  
|  |  |  |--880  
|  |  |  |  |--A1  
|  |  |  |  |  |--00000001.tif  
|  |  |  |  |  |--00000002.tif  
  
... ..  
|  |  |  |  |  |--00000023.tif  
|  |  |  |  |  |--us-patent-image.xml  
|  |  |  |--881  
|  |  |  |  |--A1  
  
... ..
```

## 4.2.3 Image Page(s) .tif Files for Patent Applications Publications

The TIFF file name for each image page will be:

```
nnnnnnnn.tif
```

where nnnnnnnn is an eight-character field containing the page number, right-aligned with leading zeros. The page number represents the sequence of the image page within the document.

### 4.3. TIFF Header Contents

The TIFF header of each page image contains standard TIFF header tags and the following tags derived from WIPO Standard ST.35.

Tags 269, 306, and 999 have been modified from the original in ST.35. Tag 50560 has been added to accommodate content type. Tag 274 will contain a constant "1" to identify that each image will be a portrait page and tag 50561 has been added to accommodate the actual rotation codes of each U.S. patent image.

| ID         | Meaning of item            | Data type | Length    | Value or pointer | Remarks                                                                                                                                                                                                                                                                                                                                      |
|------------|----------------------------|-----------|-----------|------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 254        | New subfile type           | 4         | 1         | 0                | Indicates that it is a full resolution image. Default value 0.                                                                                                                                                                                                                                                                               |
| 255        | Old subfile type           | 3         | 1         | 1                | For compatibility reasons still available.                                                                                                                                                                                                                                                                                                   |
| 256        | Width of image             | 3         | 1         | number           | In pixels (X direction).                                                                                                                                                                                                                                                                                                                     |
| 257        | Length of image            | 3         | 1         | number           | In pixels (Y direction).                                                                                                                                                                                                                                                                                                                     |
| 258        | Bits per sample            | 3         | 1         | 1                | Black and white, 1 bit per sample.                                                                                                                                                                                                                                                                                                           |
| 259        | Compression method         | 3         | 1         | 4                | ITU-T (CCITT) Fax Group 4.                                                                                                                                                                                                                                                                                                                   |
| 262        | Photometric interpretation | 3         | 1         | 0                | Minimum value (0) is white, maximum value (1) is black.                                                                                                                                                                                                                                                                                      |
| 266        | Fill order                 | 3         | 1         | 1                | Left to right.                                                                                                                                                                                                                                                                                                                               |
| <b>269</b> | <b>Document name</b>       | <b>2</b>  | <b>25</b> | <b>xx</b>        | <b>xx is a pointer to the full document number (based on WIPO Standard ST.14) as follows: Publishing office country code (2 positions); Document number (12 positions, right justified, left padded with zeros); Kind code (two positions); Date (eight positions, CCYYMMDD). The last position of this field will contain a null value.</b> |
| 270        | Image description          | 2         | 9         | xx               | xx is a pointer to the image identification, which consists of a page number (4 positions) and a frame number (4 positions) + 1 end byte.                                                                                                                                                                                                    |
| 273        | Strip offset               | 4         | 1         | xx               | xx is a pointer to the start of the image data belonging to this directory.                                                                                                                                                                                                                                                                  |
| <b>274</b> | <b>Orientation</b>         | <b>3</b>  | <b>1</b>  | <b>0</b>         | <b>Rotation or orientation of image: A constant 1 will be present denoting a Portrait image.</b>                                                                                                                                                                                                                                             |
| 277        | Samples Per Pixel          | 3         | 1         | 1                | Black and white.                                                                                                                                                                                                                                                                                                                             |
| 278        | Rows per strip             | 4         | 1         | number           | Number of rows (equal to tag 257, height in pixels).                                                                                                                                                                                                                                                                                         |
| 279        | Strip byte count           | 4         | 1         | number           | Number of bytes of image data in uncompressed form.                                                                                                                                                                                                                                                                                          |
|            |                            |           |           |                  |                                                                                                                                                                                                                                                                                                                                              |
| 280        | Min sample value           | 3         | 1         | 0                |                                                                                                                                                                                                                                                                                                                                              |
| 281        | Max sample value           | 3         | 1         | 1                |                                                                                                                                                                                                                                                                                                                                              |
| 282        | X resolution               | 5         | 1         | xx               | xx is a pointer to the field containing the numerator of the resolution in pixels in x direction, which is 4 bytes long. The value of this field is 300. The denominator follows this field immediately and is also 4 bytes long. The value of this field is 1. The result is a value of 300 DPI in x direction.                             |
| 283        | Y resolution               | 5         | 1         | xx               | Resolution in y direction, see tag 282 for exp.                                                                                                                                                                                                                                                                                              |
|            |                            |           |           | <b>Page 8</b>    |                                                                                                                                                                                                                                                                                                                                              |

| ID                | Meaning of item              | Data type | Length    | Value or pointer | Remarks                                                                                                                                     |
|-------------------|------------------------------|-----------|-----------|------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| 293               | Group 4 options              | 4         | 1         | 0                | Compressed in ITU-T (CCITT) Gr 4 format.                                                                                                    |
| 296               | Resolution unit              | 3         | 1         | 2                | Inches.                                                                                                                                     |
| <b>306</b>        | <b>Date time</b>             | <b>2</b>  | <b>20</b> | <b>xx</b>        | <b>xx is a pointer to the field containing the Date (YYYY:MM:DD) and the Time (HH:MM:SS). This is the creation date of the TIFF header.</b> |
| 999               | Miscellaneous                | 2         | 253       | xx               | Private field. By default, this field is blank.                                                                                             |
| <b>5056<br/>0</b> | <b>Original content type</b> | <b>3</b>  | <b>1</b>  | <b>0</b>         | <b>0 = text or black &amp; white drawing (default);<br/>1 = grayscale drawing or photograph; 2 = color drawing or photograph</b>            |
| <b>5056<br/>1</b> | <b>Rotation Code</b>         | <b>3</b>  | <b>1</b>  | <b>0</b>         | <b>Rotation or orientation of image:<br/>1 = portrait, 6 = landscape</b>                                                                    |

#### 4.4. Metadata File DTD

For each document there will have a metadata file that is an instance of the following document type definition. The file name of the metadata file for each document will be us-patent-image.xml.

```
<!--Document Type Definition for metadata to accompany facsimile images of United States patents.
Reference this DTD as PUBLIC "-//USPTO//DTD us-patent-image v1.0 2002-06-04//EN"
Alias: Yellow Book 2 (YB2)
Contact: Ed Johnson
Information Products Division
U.S. Patent and Trademark Office
600 Dulany Street, MDW04C03
Alexandria, VA 22314
vox: 703-306-2621
fax: 703-306-2737
ed.johnson@uspto.gov
```

```
***** Revision History *****
2012-08-27 Ed Johnson
```

Identify Kind Codes "C", "F", "J", "K" and "O" as representing the following Certificates"

Cn - Reexamination Certificate issued on or after *January 2, 2001*  
NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

There are three (3) types of Reexamination Certificates for kind code "C":  
(1). Ex Parte Reexamination Certificate under 35 U.S.C. 307  
(2). Ex Parte Reexamination Certificate under 35 U.S.C. 257  
(3). Inter Partes Reexamination Certificate under 35 U.S.C. 316

Fn - Supplemental Examination Certificate published after September 16, 2012  
NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

Jn - Post Grant Review Certificate published after September 16, 2012  
NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

Kn - Inter Partes Review Certificate published after September 16, 2012  
NOTE: "n" represents a value 1 through 9 denoting the publication  
sequence.  
On - Derivation Certificate published after March 16, 2013

2003-06-10 Barry Frank  
. Changed all references of element name "drawup" to "scan-date".  
. Changed all references of element name "withdrawn-flag" to "withdrawn-  
indicator".  
. Changed all references of element name "start" to "begin". Also changed  
comments referring to start  
.. to refer to begin.

2003-03-28 Barry Frank  
. Added bib-pages?,abstract-pages?,drawings-pages?,description-pages?,claims-  
pages? to  
.. the reexamination-certificate element.  
. Removed the ? from the related-document element in the certificate-of-  
correction  
.. and reexamination-certificate elements. (A related document must be  
present)

2002-06-18 Bruce B. Cox  
. Final version 1. Added withdrawn as valid status type.

2002-06-04 Bruce B. Cox  
. Final draft of version 1. Eliminated page metadata content model and revised  
document metadata  
.. content model. All page-specific information now in TIFF header, for a  
description of which, see YB2  
.. specification.

2002-05-10 First public draft.  
\*\*\*\*\* End Revision History \*\*\*\*\*  
-->

```
<!ELEMENT us-patent-image (patent-metadata?,certificate-of-correction*,  
reexamination-certificate*) >
```

```
<!ATTLIST us-patent-image  
file CDATA #REQUIRED  
file-type (tiff) #FIXED "tiff"  
date-produced CDATA #REQUIRED  
lang CDATA #REQUIRED  
dtd-version CDATA #IMPLIED  
status CDATA #IMPLIED  
country CDATA #FIXED "us" >
```

<!--For both US Patent Grants and US Patent Application Publications. The data-  
capture contractor will use patent-metadata for all deliverables (grants,  
applications, certificates of correction, and reexamination certificates).  
Dissemination products, however, will use patent-metadata, certificate-of-  
correction, and reexamination-certificate appropriately.-->

```
<!ELEMENT patent-metadata (full-document-number,document-id,page-count,scan-  
date, record-status,related-document?,withdrawn-indicator?,missing-pages?,bib-  
pages?,abstract-pages?,drawings-pages?,description-pages?,claims  
pages?,certificate-of-correction-pages?,reexamination-pages?) >
```

<!--Begin and End indicate the first and last pages of just this one certificate  
of correction relative to the entire document-->

```
<!ELEMENT certificate-of-correction (document-id,page-count,scan-date,record-  
status, related-document,missing-pages?,begin,end) >
```



```

<!--Begin and End indicate the first and last pages of just this one
reexamination certificate relative to the entire document-->
<!ELEMENT reexamination-certificate (document-id,page-count,scan-date,record-
status,
        related-document,missing-pages?,begin,end,bib-pages?,abstract-
pages?,drawings-pages?,description-pages?,
        claims-pages?) >

<!--The complete document identification, arranged for display, as in ST.14-->
<!ELEMENT full-document-number (#PCDATA) >

<!--Document identification refers to patents and patent applications only. See
WIPO ST.14-->

<!ELEMENT document-id (country,doc-number,kind,name?,date?) >

<!ATTLIST document-id
        lang CDATA #IMPLIED >

<!--Total number of image pages in the document.-->
<!ELEMENT page-count (#PCDATA) >

<!--Date that page image(s) were created.-->
<!ELEMENT scan-date (date) >

<!--New = page images of a new publication
Rescan = some or all of the image pages have been replaced with corrected images,
or addition of missing pages
Delete = all images of the referenced document should be deleted-->

<!ELEMENT record-status EMPTY >
<!ATTLIST record-status
        value (new | rescan | retro | delete | withdrawn) #REQUIRED >

<!--If the document is a reissue patent, this is the number of the original
document. If the document is a certificate of correction, this is the number of
the corrected document.-->
<!ELEMENT related-document (doc-number) >

<!--Indicates that the document has been withdrawn.-->
<!ELEMENT withdrawn-indicator EMPTY >

<!--Contains a list of missing pages, comma separated. If the element is present
but no page numbers are present, there are pages known to be missing, but the
page numbers are unknown.-->
<!ELEMENT missing-pages (#PCDATA) >

<!--The first (begin) and last (end) pages with bibliographic information.
Normally, begin will always = 1.-->
<!ELEMENT bib-pages (begin,end) >

<!--The first and last pages on which the abstract appears. For US documents, the
abstract normally begins on page 1.-->
<!ELEMENT abstract-pages (begin,end) >

```

<!--The first and last page numbers of the drawing pages. In US documents, drawings normally follow the abstract and precede the description. Drawing pages do not overlap with the preceding or following subdocuments.-->  
<!ELEMENT drawings-pages (begin,end) >

<!--The first and last pages of the description. The last page of the description might be the same as the first page of the claims. Sequence listings are normally between the description and the claims.-->  
<!ELEMENT description-pages (begin,end) >

<!--The first and last page of the claims. The first page of claims might be the same as the last page of the description. Sequence listings are usually between the description and the claims.-->  
<!ELEMENT claims-pages (begin,end) >

<!--The first page of the first certificate of correction and the last page of the last certificate of correction.-->  
<!ELEMENT certificate-of-correction-pages (begin,end) >

<!--The first page of the first reexamination certificate and the last page of the last reexamination certificate.-->  
<!ELEMENT reexamination-pages (begin,end) >

<!--First image page on which there is any part of the subdocument in question.-->  
>  
<!ELEMENT begin (#PCDATA) >

<!--Last image page on which there is any part of the document in question.-->  
<!ELEMENT end (#PCDATA) >

<!--Country: use ST.3 country code, e.g. DE, FR, GB, NL, etc. Also includes EP, WO, and other regional authorities.  
ST.32 name: B190; B330-->  
<!ELEMENT country (#PCDATA) >

<!--The number of the referenced patent (or application) document.  
ST.32 name: B110; B210; B310-->  
<!ELEMENT doc-number (#PCDATA) >

<!--The doc-number for Reexamination Certificates will contain a leading "C" following by the original patent number of the patent being reexamined-->

<!--Document kind code; e.g.  
A1 - Utility Patent Grant issued prior to January 2, 2001  
A1 - Utility Patent Application published on or after January 2, 2001  
A9 - Correction published Utility Patent Application  
A2 - *Second or subsequent publication of a Utility Patent Application*  
B1 - Utility Patent Grant (no published application) issued on or after January 2, 2001  
B2 - Utility Patent Grant (with a published application) issued on or after January 2, 2001  
Bn - Reexamination Certificate issued prior to *January 2, 2001*  
NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

Cn - Reexamination Certificate issued on or after *January 2, 2001*  
NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

There are three (3) types of Reexamination Certificates for kind code "C":

- (1). Ex Parte Reexamination Certificate under 35 U.S.C. 307
- (2). Ex Parte Reexamination Certificate under 35 U.S.C. 257
- (3). Inter Partes Reexamination Certificate under 35 U.S.C. 316

E1 - Reissue Patent

Fn - Reexamination Certificate of a Reissue issued on or prior to January 12, 2010

NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

Fn - Supplemental Examination Certificate published after September 16, 2012  
NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

H1 - Statutory Invention Registration (SIR) Patent Documents

Note: SIR documents began with the December 3, 1985 issue

I1 - "X" Patents issued from July 31, 1790 prior to July 3, 1836

I2 - "X" Reissue Patents issued from July 7, 1836 to June 13, 1848

I3 - Additional Improvements - Patents issued between 1838 and 1861

I4 - Defensive Publication - Documents issued from November 5, 1968 through May 5, 1987

I5 - Trial Voluntary Protest Program (TVPP) Patent Documents

Jn - Post Grant Review Certificate published after September 16, 2012

NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

Kn - Inter Partes Review Certificate published after September 16, 2012

NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

NP - Non-Patent Literature

On - Derivation Certificate published after March 16, 2013

NOTE: "n" represents a value 1 through 9 denoting the publication sequence.

P1 - Plant Patent Grant issued prior to January 2, 2001

P1 - Plant Patent Application published on or after January 2, 2001

P2 - Plant Patent Grant (no published application) issued on or after January 2, 2001

P3 - Plant Patent Grant (with a published application) issued on or after January 2, 2001

P4 - Second or subsequent publication of a Plant Patent Application

P9 - Correction publication of a Plant Patent Application

S1 - Design Patent

X6 - Certificates of Correction -->

<!ELEMENT kind (#PCDATA) >

<!ELEMENT name (#PCDATA) >

<!ATTLIST name  
                  name-type (legal | natural) #IMPLIED >

<!--Format: YYYYMMDD-->

<!ELEMENT date (#PCDATA) >

## 4.5. Metadata File Content Examples

### 4.5.1. Patent Grant Metadata

```
<?xml version="1.0"?>
<!DOCTYPE us-patent-image PUBLIC "-//USPTO//DTD us-patent-image v1.0 2002-06-04//EN"
"..\\..\\..\\us-patent-image-010.dtd">
<us-patent-image file="07401653" date-produced="20080702" lang="EN">
<patent-metadata>
  <full-document-number>07401653</full-document-number>
  <document-id>
    <country>US</country>
    <doc-number>07401653</doc-number>
    <kind>B2</kind>
    <date>20080722</date>
  </document-id>
  <page-count>4</page-count>
  <scan-date>
    <date>20080702</date>
  </scan-date>
  <record-status value="new" />
  <bib-pages>
    <begin>1</begin>
    <end>1</end>
  </bib-pages>
  <abstract-pages>
    <begin>1</begin>
    <end>1</end>
  </abstract-pages>
  <drawings-pages>
    <begin>0</begin>
    <end>0</end>
  </drawings-pages>
  <description-pages>
    <begin>2</begin>
    <end>3</end>
  </description-pages>
  <claims-pages>
    <begin>3</begin>
    <end>4</end>
  </claims-pages>
</patent-metadata>
</us-patent-image>
```

**4.5.1.a.** The data content of the <full-document-number> and the <doc-number> xml tags will contain an 8-position patent number. Reference Table 1a - U.S. Patent Grant Patent Numbers

**4.5.1.b.** The data content of the <country> xml tag within the <document-id> will contain “US” identifying the United States as the publishing country

**4.5.1.c.** The data content of the <kind> xml tag within the <document-id> will contain a 2-position kind code. Reference Table 2 – U.S. Patent Grants and Patent Published Applications – Kind Codes

**4.5.1.d.** The data content of the <date> xml tag within the <document-id> will contain the date, in yyyyymmdd format, of the patent grant.

**4.5.1.e.** The data content of the <page-count> xml tag will contain the page total of the patent grant.

**4.5.1.f.** The data content of the <date> xml tag within the <scan-date> will contain the date, in yyyyymmdd format, that the TIFF image pages were captured.

**4.5.1.g.** The data content of the <record-status value=" " /> xml tag will contain one of the following values:

- “new” for a new patent document or certificate of correction
- “rescan” for a rescanned patent document
- “retro” for a reexamined patent document

**4.5.1.h.** The data content of each <begin> xml tag and <end> xml tag will identify the beginning pages for the appropriate section of the patent grant.

#### **4.5.2. Patent Published Application Metadata**

```
<?xml version="1.0"?>
<!DOCTYPE us-patent-image PUBLIC "-//USPTO//DTD us-patent-image v1.0 2002-06-04//EN"
"..\\.\\.\\.us-patent-image-010.dtd">
<us-patent-image file="11626986" date-produced="20080715" lang="EN">
<patent-metadata>
<full-document-number>US20080179035A1</full-document-number>
<document-id>
<country>US</country>
<doc-number>20080179035</doc-number>
<kind>A1</kind>
<date>20080731</date>
</document-id>
<page-count>4</page-count>
<scan-date>
<date>20080715</date>
</scan-date>
<record-status value="new" />
<bib-pages>
<begin>1</begin>
<end>1</end>
</bib-pages>
<abstract-pages>
<begin>1</begin>
<end>1</end>
</abstract-pages>
```

```
<drawings-pages>
<begin>2</begin>
<end>2</end>
</drawings-pages>
<description-pages>
<begin>3</begin>
<end>4</end>
</description-pages>
<claims-pages>
<begin>4</begin>
<end>4</end>
</claims-pages>
</patent-metadata>
</us-patent-image>
```

**4.5.2.a.** The data content of the <full-document-number> xml tag will contain the following. Reference Table 1b - U.S. Patent Application Publication Number

- A 2-position country code “US”
- A 4-position numeric year of publication
- A 7-position numeric sequence number, right justified with leading zeros
- A 2-position kind code - Reference Table 2 – U.S. Patent Grants and Patent Published Applications – Kind Codes

**4.5.2.b.** The data content of the <country> xml tag within the <document-id> will contain “US” identifying the United States as the publishing country.

**4.5.2.c.** The data content of the <doc-number> xml tag within the <document-id> will contain a 4-position numeric year of publication and a 7-position numeric sequence number, right justified with leading zeros

**4.5.2.d.** The data content of the <kind> xml tag within the <document-id> will contain a 2-position kind code - Reference Table 2 – U.S. Patent Grants and Patent Published Applications – Kind Codes

**4.5.2.e.** The data content of the <date> xml tag within the <document-id> will contain the date, in yyyyymmdd format, of the patent publication application date

**4.5.2.f.** The data content of the <page-count> xml tag will contain the page total of the patent published application

**4.5.2.g.** The data content of the <date> xml tag within the <scan-date> will contain the date, in yyyyymmdd format, that the TIFF image pages were captured.

**4.5.2.h.** The data content of the <record-status value=" " /> xml tag will contain one of the following values:

- “new” for a new patent document or certificate of correction
- “rescan” for a rescanned patent document
- “retro” for a reexamined patent document

**4.5.2.i.** The data content of each <begin> xml tag and <end> xml tag will identify the beginning pages for the appropriate section of the patent grant.

## Table 1a - U.S. Patent Grant Patent Numbers

**Utility Patents** – Positions 1-8 – 8 numeric positions, right justified, with a leading zero.

**Design Patents** – Position 1 – A constant “D” identifying the granted document as a Design Patent.  
Positions 2-8 – 7 numeric positions, right justified, with a leading zero.

**Certificates** - Patents Numbers for the following Certificates will contain the 8-position number of the original published document:

**Supplemental Examination Certificate** published after September 16, 2012

**Post Grant Review Certificate** published after September 16, 2012

**Inter Partes Review Certificate** published after September 16, 2012

**Derivation Certificate** published after March 16, 2013

### Reexamination Certificates:

- Utility “B” followed by 7 numeric positions, issued on prior to January 2, 2001
- Design “BD” followed by 6 numeric positions, issued prior to January 2, 2001
- Plant “BP” followed by 6 numeric positions, right justified, with a leading zeros, issued prior to January 2, 2001
- Reissue “BRE” followed by 5 numeric positions, issued prior to January 2, 2001
  
- Utility “C” followed by 7 numeric positions, issued on or after January 2, 2001
- Design “CD” followed by 6 numeric positions, issued on or after January 2, 2001
- Plant “CP” followed by 6 numeric positions, right justified, with a leading zeros, issued on or after January 2, 2001
- Reissue “CRE” followed by 5 numeric positions, issued on or after January 2, 2001

**Plant Patents** – Positions 1-2 – A constant “PP” identifying the granted document as a Plant Patent.  
Positions 3-8 – 6 numeric positions, right justified, with a leading zero.

**Reissue Patents** – “RE” followed by 6 numeric positions, with leading zeros.

**SIR Patents** – SIR Utility – “H” followed by 7 numeric positions.

## Table 1a - U.S. Patent Grant Patent Numbers (continued)

**X-Series Patents** – Utility Patents issued from July 31, 1790 to July 3, 1836. They were not originally numbered, but have since been assigned numbers in the sequence in which they were issued.

- Utility “X” Patent Numbers will contain “X” followed by 7 numeric positions, with a leading zero

**RX-Series Patents** – Reissue Patents issued from July 7, 1836 to June 13, 1848 to replace the X parent patents destroyed in the December 1836 fire.

- Reissued X Patent Numbers will contain “RX” followed by 6 numeric positions, with a leading zero.

**Fraction Patents** – 89 early patents exist that are considered fraction “X” patents because the patent number ends with a fraction. To accommodate this fraction the last character of the patent number contains an upper case alpha character that is equated to the appropriate fraction:

- Utility Fraction Patent Numbers will contain “X” followed by 6 numeric positions, with leading zeros. The 8<sup>th</sup> position of a utility fraction patent number will contain an alpha character as defined below
- Design Fraction Patent Numbers will contain “D” followed by 6 numeric positions, with leading zeros. The 8<sup>th</sup> position of a design fraction patent number will contain an alpha character as defined below
- Reissued Fraction Patent Numbers will contain “RE” followed by 5 numeric positions, with leading zeros. The 8<sup>th</sup> position of a reissue fraction patent number will contain an alpha character as defined below

|          |           |
|----------|-----------|
| A - 1/16 | I – 9/16  |
| B – 1/8  | J – 5/8   |
| C – 3/16 | K – 11/16 |
| D – 1/4  | L – 3/4   |
| E – 5/16 | M – 13/16 |
| F – 3/8  | N – 7/8   |
| G – 7/16 | O – 15/16 |
| H – 1/2  |           |

## Table 1b - U.S. Patent Application Publication Number

A 2-position country code “US”  
A 4-position numeric year of publication  
A 7-position numeric sequence number, right justified with leading zeros  
A 2-position kind code - Reference Table 2 – U.S. Patent Grants and Patent Published Applications – Kind Codes



**Table 2 - U.S. Patent Grants and Patent Published Applications –  
Kind Codes**

Note: The following 2-position kind codes will be present in the XML <kind> tags of Yellow Book. These 2-positions kind codes will also be present on the printed documents with the following exceptions: Reissues will contain a single position “E”, SIR documents will contain a single position “H”, and Designs will contain a single position “S”.

- A1** - Utility Patent Grant issued prior to January 2, 2001
- A1** - Utility Patent Application published on or after January 2, 2001
- A9** - Correction published Utility Patent Application
- A2** - *Second or subsequent publication of a Utility Patent Application*
- B1** - Utility Patent Grant (no published application) issued on or after January 2, 2001
- B2** - Utility Patent Grant (with a published application) issued on or after January 2, 2001
- Bn** - Reexamination Certificate issued prior to *January 2, 2001*  
NOTE: “n” represents a value 1 through 9 denoting the publication sequence.
- Cn** - Reexamination Certificate issued on or after *January 2, 2001*  
NOTE: “n” represents a value 1 through 9 denoting the publication sequence.

There are three (3) types of Reexamination Certificates for kind code “C”:

- (1). Ex Parte Reexamination Certificate under 35 U.S.C. 307
- (2). Ex Parte Reexamination Certificate under 35 U.S.C. 257
- (3). Inter Partes Reexamination Certificate under 35 U.S.C. 316

- E1** - Reissue Patent
- Fn** - Reexamination Certificate of a Reissue issued on or prior to January 12, 2010  
NOTE: “n” represents a value 1 through 9 denoting the publication sequence.
- Fn** – Supplemental Examination Certificate published after September 16, 2012  
NOTE: “n” represents a value 1 through 9 denoting the publication sequence.

- H1** - Statutory Invention Registration (SIR) Patent Documents  
Note: SIR documents began with the December 3, 1985 issue
- I1** - “X” Patents issued from July 31, 1790 prior to July 3, 1836
- I2** - “X” Reissue Patents issued from July 7, 1836 to June 13, 1848
- I3** - Additional Improvements – Patents issued between 1838 and 1861
- I4** - Defensive Publication – Documents issued from November 5, 1968 through May 5, 1987
- I5** - Trial Voluntary Protest Program (TVPP) Patent Documents

**Table 2 - U.S. Patent Grants and Patent Published Applications –  
Kind Codes (continued)**

**Jn** - Post Grant Review Certificate published after September 16, 2012

NOTE: “n” represents a value 1 through 9 denoting the publication sequence.

**Kn** - Inter Partes Review Certificate published after September 16, 2012

NOTE: “n” represents a value 1 through 9 denoting the publication sequence.

**NP** - Non-Patent Literature

**On** - Derivation Certificate published after March 16, 2013

NOTE: “n” represents a value 1 through 9 denoting the publication sequence.

**P1** - Plant Patent Grant issued prior to January 2, 2001

**P1** - Plant Patent Application published on or after January 2, 2001

**P2** - Plant Patent Grant (no published application) issued on or after  
January 2, 2001

**P3** - Plant Patent Grant (with a published application) issued on or after  
January 2, 2001

**P4** - Second or subsequent publication of a Plant Patent Application

**P9** - Correction publication of a Plant Patent Application

**S1** - Design Patent

**X6** - Certificates of Correction