**Yellow Book 2**

**U. S. Patent Images/TIFF**

**United States Patents**

**(Grants and Published Applications)**

**Delivered as**

**CCITT Group 4 Facsimile Images**

**Revised March 15, 2018**

**United States Patent & Trademark Office**

**Electronic Information Products Division**

|  |
| --- |
| **Update to U.S. Patent Image/TIFF (a.k.a Yellow Book 2) Table** |
| **Date** | **Sections Modified** | **Comment** |
| **March 15, 2018** | **2. Summary** | **Minor wording changes.** |
| **March 15, 2018** | **3.2. Content List File (a.k.a. TCL)** | **Updated to include examples covering 8 to 13 digits patent numbers** |
| **March 15, 2018** | **4.1.1 Directory Structure Hierarchy Patent Grant Publications** | **Updated to include examples covering 8 to 13 digits patent numbers** |
| **March 15, 2018** | **4.1.2 Directory Structure Example - Patent Grant Publications** | **Updated to include examples covering 8 to 13 digits patent numbers** |
| **March 15, 2018** | **4.3. TIFF Header Contents** | **Updated t to include examples covering 8 to 13 digits patent numbers** |
| **March 15, 2018** | **4.5. Metadata File Content Examples** | **Updated to add file name of XML document**  |
| **March 15, 2018** | **4.5.2.i.** | **Updated to include examples covering 8 to 13 digits patent numbers** |

**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 Issued Patent Grant Documents and weekly Pre-Grant Publications (Application), as well as weekly Certificates of Correction and daily certificates (Reexamination, Supplemental Examination, Patent Trial and Appeal Board) 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 Documents and Pre-Grant Published Applications) will be electronic files. Also included with each electronic file 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.

|  |
| --- |
| ***Media Series Code*** |
| A | Patent Grants prior to June 4, 2002 |
| BB | Replacement Documents captured from paper,Certificates of Correction captured from paper. |
| G | Patent Grants  |
| M orD | Replacement Documents, Certificates of Correction, Reexamination Certificates dissemination to the patent examining search system1.  |
| P | Patent Application Publications |
| R | Reexamination Certificates – Disseminated on separate weekly optical disc beginning October 4, 2011 |
| R5 | Certificates of Extension |
| T | Replacement Patent Application Publications2 |
| ZZ | Certificates of Correction captured electronically Certificates of Correction (Patent Term Adjustment)captured electronically  |

1 The Maintenance “M” or “D” file (“M” and “D” are equivalent) 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”.

2 The Maintenance “T” file for the Patent Application Publications examining search system will contain: Replacement documents with the <status> field containing “RESCAN”.

## **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 certificates of correction and patent application publications (P), and for daily publication of certificates (Reexamination, Supplemental Examination, Patent Trial and Appeal Board), 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 tab - “hex 09” and each record/document terminated by a linefeed character - “hex 0D0A”,

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

**20140923.contents:**

08839462 B2 20140923 10

08839463 B2 20140923 6

08839464 B2 20140923 41

08839465 B2 20140923 9

08839466 B2 20140923 13

09999999 B2 20140923 13

100000000 B2 20140923 13

1000000000 B2 20140923 13

10000000000 B2 20140923 13

100000000000 B2 20140923 13

1000000000000 B2 20140923 13

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

**20140909.contents:**

06255118 X6 20140909 1

06363295 X6 20140909 1

06445777 X6 20140909 1

06490822 X6 20140909 1

07178274 X6 20140909 1

09999999 X6 20140909 1

100000000 X6 20140909 1

1000000000 X6 20140909 1

10000000000 X6 20140909 1

100000000000 X6 20140909 1

1000000000000 X6 20140909 1

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

**20140902.contents:**

C5974120 C2 20140902 3

C6098203 C1 20140902 2

C6465961 C1 20140902 6

C6768999 C1 20140902 2

C6933505 C1 20140902 2

C9999999 C1 20140902 2

C100000000 C1 20140902 2

C1000000000 C1 20140902 2

C10000000000 C1 20140902 2

C100000000000 C1 20140902 2

C1000000000000 C1 20140902 2

**Example of a Content List for** **Pre-Grant Published Applications:**

|  |  |  |  |
| --- | --- | --- | --- |
| **20140904.contents:**US20140245516A1 A1 20140904 40US20140245517A1 A1 20140904 9US20140245518A1 A1 20140904 12US20140245519A1 A1 20140904 8US20140245520A1 A1 20140904 6  |  |  |  |
|  |  |  |  |

## **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 were created or modified.

# **4. Document Image Pages**

## **4.1. Directory Structure Patent Grant Documents**

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 Documents**

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

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Patent No. Position->** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Description |
| **7 digit (has leading zero)** | **YYYY-WW** |  |  |  |  |  |  |  |  |  |  |  |  |  | 4 digit Year-2 digit week number |
|  | **n** | **n** |  |  |  |  |  |  |  |  |  |  |  | Position 1 and 2 |
|  |  |  | **n** | **n** | **n** |  |  |  |  |  |  |  |  | Position 3, 4, and 5 |
|  |  |  |  |  |  | **n** | **n** | **n** |  |  |  |  |  | Position 6, 7, and 8 |
| **8 digit** | **YYYY-WW** |  |  |  |  |  |  |  |  |  |  |  |  |  | 4 digit Year-2 digit week number |
|  | **n** | **n** |  |  |  |  |  |  |  |  |  |  |  | Position 1 and 2 |
|  |  |  | **n** | **n** | **n** |  |  |  |  |  |  |  |  | Position 3, 4, and 5 |
|  |  |  |  |  |  | **n** | **n** | **n** |  |  |  |  |  | Position 6, 7, and 8 |
| **9 digit** | **YYYY-WW** |  |  |  |  |  |  |  |  |  |  |  |  |  | 4 digit Year-2 digit week number |
|  | **n** | **n** | **n** |  |  |  |  |  |  |  |  |  |  | Position 1 to 3 |
|  |  |  |  | **n** | **n** | **n** |  |  |  |  |  |  |  | Position 4 to 6 |
|  |  |  |  |  |  |  | **n** | **n** | **n** |  |  |  |  | Position 7 to 9 |
| **10 digit** | **YYYY-WW** |  |  |  |  |  |  |  |  |  |  |  |  |  | 4 digit Year-2 digit week number |
|  | **n** | **n** | **n** | **n** |  |  |  |  |  |  |  |  |  | Position 1 to 4 |
|  |  |  |  |  | **n** | **n** | **n** |  |  |  |  |  |  | Position 5 to 7 |
|  |  |  |  |  |  |  |  | **n** | **n** | **n** |  |  |  | Position 8 to 10 |
| **11 digit** | **YYYY-WW** |  |  |  |  |  |  |  |  |  |  |  |  |  | 4 digit Year-2 digit week number |
|  | **n** | **n** | **n** | **n** | **n** |  |  |  |  |  |  |  |  | Position 1 to 5 |
|  |  |  |  |  |  | **n** | **n** | **n** |  |  |  |  |  | Position 6 to 8 |
|  |  |  |  |  |  |  |  |  | **n** | **n** | **n** |  |  | Position 9 to 11 |
| **12 digit** | **YYYY-WW** |  |  |  |  |  |  |  |  |  |  |  |  |  | 4 digit Year-2 digit week number |
|  | **n** | **n** | **n** | **n** | **n** | **n** |  |  |  |  |  |  |  | Position 1 to 6 |
|  |  |  |  |  |  |  | **n** | **n** | **n** |  |  |  |  | Position 7 to 9 |
|  |  |  |  |  |  |  |  |  |  | **n** | **n** | **n** |  | Position 10 to 12 |
| **13 digit** | **YYYY-WW** |  |  |  |  |  |  |  |  |  |  |  |  |  | 4 digit Year-2 digit week number |
|  | **n** | **n** | **n** | **n** | **n** | **n** | **n** |  |  |  |  |  |  | Position 1 to 7 |
|  |  |  |  |  |  |  |  | **n** | **n** | **n** |  |  |  | Position 8 to 10 |
|  |  |  |  |  |  |  |  |  |  |  | **n** | **n** | **n** | Position 11 to 13 |

*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 multi-position patent numbers intended as follows to ensure that there are no more than 1,000 subdirectories in a directory, which pertains to the directory that holds each individual patent grant:

1) – a multi-position subdirectory identifying position-1 to position-7 depending on the type and length of the patent number(s). The leading zero will not be added once patent number reaches 10 million and thereafter for Utility Patents.

-“nn” - Two-positions numeric with leading zero for patent number in the 7 digits range- Utility Patents

-“nn” - Two-positions numeric for patent number in the 8 digits range- Utility Patents

-“nnn” - Three-positions numeric for patent number in the 9 digits range- Utility Patents

-“nnnn” - Four-positions numeric for patent number in the 10 digits range- Utility Patents

-“nnnnn” - Five-positions numeric for patent number in the 11 digits range- Utility Patents

-“nnnnnn” – Six-positions numeric for patent number in the 12 digits range- Utility Patents

-“nnnnnnn” – Seven-positions numeric for patent number in the 13 digits range- Utility Patents

-“D0” - Design Patents

-“PP” – Plant Patents

-“RE” – Reissue Patents

-“H0” - Statutory Invention Registration (SIR)

-“AI” - Additional Improvements

-“T0” - Defensive Publication

 2) – a three-position subdirectory identifying position-3, position-4 and position-5 of

 the patent number(s) for all patents except for Utility Patents. Follow the guidelines below for Utility Patents.

* a three-position subdirectory identifying position-3, position-4, position-5 of patent number(s) for an 8 digits patent number.
* a three-position subdirectory identifying position-4, position-5, position-6 of patent number(s) for a 9 digits patent number.
* a three-position subdirectory identifying position-5, position-6, position-7 of patent number(s) for a 10 digits patent number.
* a three-position subdirectory identifying position-6, position-7, position-8 of patent number(s) for an 11 digits patent number.
* a three-position subdirectory identifying position-7, position-8, position-9 of patent number(s) for a 12 digits patent number.
* a three-position subdirectory identifying position-8, position-9, position-10 of patent number(s) for a 13 digits patent number.

3) – a three-position subdirectory identifying position-6, position-7 and position-8 of the patent number(s) for all Patents except for Utility Patents. Follow the guidelines below for Utility Patents.

* a three-position subdirectory identifying position-6, position-7, position-8 of patent number(s) for a 8 digits patent number.
* a three-position subdirectory identifying position-7, position-8, position-9 of patent number(s) for a 9 digits patent number.
* a three-position subdirectory identifying position-8, position-9, position-10 of patent number(s) for a 10 digits patent number.
* a three-position subdirectory identifying position-9, position-10, position-11 of patent number(s) for an 11 digits patent number.
* a three-position subdirectory identifying position-10, position-11, position-12 of patent number(s) for a 12 digits patent number.
* a three-position subdirectory identifying position-11, position-12, position-13 of patent number(s) for a 13 digits patent number.

**4.1.2 Directory Structure Example - Patent Grant Documents**

A root directory listing for patent grants issued the 5th week of 2002,

Issue date - 20020129

 G00001.tid

 20020129.contents

 2002-05

A directory listing for new patent grants issued in the 5th week of 2002, showing the subdirectory for document 6,342,021 follows:

 2002-05

 |-06

 | |--245

 | | |--001

 | | |--002

 … … …

 | |--342

 | | |--021

 | | | |--00000001.tif

 | | | |--00000002.tif

 … … … … …

 | | | |--00003999.tif

 | | | |--us-patent-image.xml

 | | |--022

 … … … …

A directory listing for new patent grants issued in the 5th week of 2018, showing the subdirectory for document 16,342,021 (8 digits) follows:

 2018-05

 |-16

 | |--245

 | | |--001

 | | |--002

 … … …

 | |--342

 | | |--021

 | | | |--00000001.tif

 | | | |--00000002.tif

 … … … … …

 | | | |--00003999.tif

 | | | |--us-patent-image.xml

 | | |--022

 … … … …

A directory listing for new patent grants issued in the 5th week of 2050, showing the subdirectory for document 116,342,021 (9 digits) follows:

 2050-05

 |-116

 | |--245

 | | |--001

 | | |--002

 … … …

 | |--342

 | | |--021

 | | | |--00000001.tif

 | | | |--00000002.tif

 … … … … …

 | | | |--00003999.tif

 | | | |--us-patent-image.xml

 | | |--022

 … … … …

A directory listing for new patent grants issued in the 5th week of 2050, showing the subdirectory for document 1,116,342,021 (10 digits) follows:

 2050-05

 |-1116

 | |--245

 | | |--001

 | | |--002

 … … …

 | |--342

 | | |--021

 | | | |--00000001.tif

 | | | |--00000002.tif

 … … … … …

 | | | |--00003999.tif

 | | | |--us-patent-image.xml

 | | |--022

 … … … …

A directory listing for new patent grants issued in the 5th week of 2070, showing the subdirectory for document 11,116,342,021 (11 digits) follows:

 2070-05

 |-11116

 | |--245

 | | |--001

 | | |--002

 … … …

 | |--342

 | | |--021

 | | | |--00000001.tif

 | | | |--00000002.tif

 … … … … …

 | | | |--00003999.tif

 | | | |--us-patent-image.xml

 | | |--022

 … … … …

A directory listing for new patent grants issued in the 5th week of 2090, showing the subdirectory for document 111,116,342,021 (12 digits) follows:

 2090-05

 |-111116

 | |--245

 | | |--001

 | | |--002

 … … …

 | |--342

 | | |--021

 | | | |--00000001.tif

 | | | |--00000002.tif

 … … … … …

 | | | |--00003999.tif

 | | | |--us-patent-image.xml

 | | |--022

 … … … …

A directory listing for new patent grants published in the 5th week of 2100, showing the subdirectory for document 1,111,116,342,021 (13 digits) follows:

 2100-05

 |-1111116

 | |--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 Documents**

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 Pre-Grant Published Applications**

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

## **4.2.1 Directory Structure Hierarchy Pre-Grant Published Applications**

The hierarchy of the directory structure containing pre-grant published applications 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 – Pre-Grant Published Applications**

A root directory listing for pre-grant published applications published the 3rd week of 2002,

Issue date - 20020117

 P12345.tid

 20020117.contents

 2002-03

A directory listing for new patent published applications published in the 3rd 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 Pre-Grant Published Applications**

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 orpointer | 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. |
| **269** | **Document name** | **2** | **25** | **xx** | **For Patent Number up to 9,999,999: (This applies to Utility, Design, Plant, and Reissues) Please see next row for instruction on 10 millionth utility patent number and greater.** **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.** |
| **269** | **Document name** | **2** | **27** | **xx** | **For Utility Patent Number 10,000,000 and greater:****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 (14 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.** |
| 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. |
| **274** | **Orientation** | **3** | **1** | **0** | **Rotation or orientation of image:** **A constant 1 will be present denoting a** **Portrait image.** |
| 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. |
|  |  |  |  | Page 8 |  |
| 283 | Y resolution | 5 | 1 | xx | Resolution in y direction, see tag 282 for exp. |
| 293 | Group 4 options  | 4 | 1 | 0 | Compressed in ITU-T (CCITT) Gr 4 format. |
| 296 | Resolution unit | 3 | 1 | 2 | Inches. |
| **306** | **Date time** | **2** | **20**  | **xx** | **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.** |
| 999 | Miscellaneous | 2 | 253 | xx | Private field. By default, this field is blank. |
| **50560** | **Original content type** | **3** | **1** | **0** | **0 = text or black & white drawing (default); 1 = grayscale drawing or photograph; 2 = color drawing or photograph** |
| **50561** | **Rotation Code** | **3** | **1** | **0** | **Rotation or orientation of image:** **1 = portrait, 6 = landscape** |

## **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: Narith Tith

Enterprise Data Architecture Division

U.S. Patent and Trademark Office

600 Dulany Street, MDW 5C87

Alexandria, VA 22314

Phone: 571-272-5458

narith.tith@uspto.gov

\*\*\*\*\*\*\*\* Revision History \*\*\*\*\*\*\*\*

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 Kind codes changed effective 2001-01-02 to accommodate pre-grant publication status.

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

A1 - Utility Patent Application published on or after January 2, 2001.

A2 - Second or subsequent publication of a Utility Patent Application.

A9 - Corrected published Utility Patent Application.

Bn - Reexamination Certificate issued prior to January 2, 2001.

B1 - Utility Patent (no pre-grant publication) issued on or after January 2, 2001.

B2 - Utility Patent (with pre-grant publication) issued on or after January 2, 2001.

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

E1 - Reissue Patent.

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

Fn – Supplemental Examination Certificate published after September 16, 2012

H1 - Statutory Invention Registration (SIR) Patent Documents.

*SIR documents began with the December 3, 1985 issue.*

*I1 - “X” Patents issued from July 31, 1790 prior to July 3, 1836.*

*I2 - Reissued “X” Patents that issued on or before 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.*

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

*NP - Non-Patent Literature.*

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

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

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

*P2 - Plant Patent (no pre-grant publication) issued on or after January 2, 2001.*

*P3 - Plant Patent (with pre-grant publication) 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. ST.32 name: B130-->*

#### Note: Reference Table 2 - U.S. Patent Grants and Patent Published Applications – *Kind Codes, in the Yellow Book 2 documentation, for a complete list of kind codes that took effect January 2, 2001.*

<!ELEMENT kind (#PCDATA) >

<!ELEMENT name (#PCDATA) >

<!ATTLIST name

 name-type (legal | natural) #IMPLIED >

<!--Format: YYYYMMDD-->

<!ELEMENT date (#PCDATA) >

**4.5. Metadata File Content Examples (name of XML document for PGPUB and Grant: us-patent-image.xml)**

**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 yyyymmdd 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 yyyymmdd 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. Pre-Grant 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. Pre-Grant Published Application 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 yyyymmdd 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 yyyymmdd 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 (s) for patent numbers less than 8 digits.

– Positions 1-8 – 8 numeric positions, for 8 digits patent numbers.

– Positions 1-9 – 9 numeric positions, for 9 digits patent numbers.

– Positions 1-10 – 10 numeric positions, for 10 digits patent numbers.

– Positions 1-11 – 11 numeric positions, for 11 digits patent numbers.

– Positions 1-12 – 12 numeric positions, for 12 digits patent numbers.

– Positions 1-13 – 13 numeric positions, for 13 digits patent numbers.

**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.

**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.

**Reexamination Certificates:**

**-** Utility - “B” followed by 7 numeric positions, right justified, with leading zeros, issued on prior to January 2, 2001

**-** Design - “BD” followed by 6 numeric positions, right justified, with leading zeros, issued prior to January 2, 2001

**-** Plant - “BP” followed by 6 numeric positions, right justified, with leading zeros,

 issued prior to January 2, 2001

**-** Reissue - “BRE” followed by 5 numeric positions, right justified, with leading zeros, issued prior to January 2, 2001

**-** Utility

- “C” followed by 7 numeric positions, for 7 digits patent numbers , issued on or after January 2, 2001

- “C” followed by 8 numeric positions, for 8 digits patent numbers

- “C” followed by 9 numeric positions, for 9 digits patent numbers

- “C” followed by 10 numeric positions, for 10 digits patent numbers

- “C” followed by 11 numeric positions, for 11 digits patent numbers

- “C” followed by 12 numeric positions, for 12 digits patent numbers

- “C” followed by 13 numeric positions, for 13 digits patent numbers

**-** Design - “CD” followed by 6 numeric positions, right justified, with leading zeros, issued on or after January 2, 2001

**-** Plant - “CP” followed by 6 numeric positions, right justified, with leading zeros,

 issued on or after January 2, 2001

**-** Reissue - “CRE” followed by 5 numeric positions, right justified, with leading zeros, issued on or after January 2, 2001

**Supplemental Examination Certificates:**

 Utility

- “F” followed by 7 numeric positions, for 7 digits patent numbers issued on or after September 16, 2012

- “F” followed by 8 numeric positions, for 8 digits patent numbers

- “F” followed by 9 numeric positions, for 9 digits patent numbers

- “F” followed by 10 numeric positions, for 10 digits patent numbers

- “F” followed by 11 numeric positions, for 11 digits patent numbers

- “F” followed by 12 numeric positions, for 12 digits patent numbers

- “F” followed by 13 numeric positions, for 13 digits patent numbers

 Design - “FD” followed by 6 numeric positions, right justified, with leading zeros, issued on or after September 16, 2012

 Plant - “FP” followed by 6 numeric positions, right justified, with leading zeros, issued on or after September 16, 2012

 Reissue - FRE” followed by 5 numeric positions, right justified, with leading zeros, issued on or after September 16, 2012

**Patent Trial and Appeal Board (PTAB) Certificates**

**Post Grant Review Certificates:**

 Utility

- “J” followed by 7 numeric positions, for 7 digits patent numbers issued on or after September 16, 2012

- “J” followed by 8 numeric positions, for 8 digits patent numbers

- “J” followed by 9 numeric positions, for 9 digits patent numbers

- “J” followed by 10 numeric positions, for 10 digits patent numbers

- “J” followed by 11 numeric positions, for 11 digits patent numbers

- “J” followed by 12 numeric positions, for 12 digits patent numbers

- “J” followed by 13 numeric positions, for 13 digits patent numbers

 Design - “JD” followed by 6 numeric positions, right justified, issued on or after September 16, 2012

 Plant - “JPP” followed by 5 numeric positions issued on or after September 16, 2012

 Reissue – “JRE” followed by 5 numeric positions issued on or after September 16, 2012

**Inter Partes Review Certificates:**

 Utility

- “K” followed by 7 numeric positions, for 7 digits patent numbers issued on or after September 16, 2012

- “K” followed by 8 numeric positions, for 8 digits patent numbers

- “K” followed by 9 numeric positions, for 9 digits patent numbers

- “K” followed by 10 numeric positions, for 10 digits patent numbers

- “K” followed by 11 numeric positions, for 11 digits patent numbers

- “K” followed by 12 numeric positions, for 12 digits patent numbers

- “K” followed by 13 numeric positions, for 13 digits patent numbers

 Design - “KD” followed by 6 numeric positions, issued on or after September 16, 2012

 Plant - “KPP” followed by 6 numeric positions issued on or after September 16, 2012

 Reissue - KRE” followed by 5 numeric positions issued on or after September 16, 2012

**Derivation Certificates:**

 Utility

- “O” followed by 7 numeric positions, for 7 digits patent numbers issued on or after March 16, 2013

- “O” followed by 8 numeric positions, for 8 digits patent numbers

- “O” followed by 9 numeric positions, for 9 digits patent numbers

- “O” followed by 10 numeric positions, for 10 digits patent numbers

- “O” followed by 11 numeric positions, for 11 digits patent numbers

- “O” followed by 12 numeric positions, for 12 digits patent numbers

- “O” followed by 13 numeric positions, for 13 digits patent numbers

 Design - “OD” followed by 6 numeric positions issued on or after March 16, 2013

 Plant - “OPP” followed by 6 numeric positions issued on or after March 16, 2013

 Reissue – “ORE” followed by 5 numeric positions issued on or after March 16, 2013

**Reissue Patents**

– “RE” followed by 6 numeric positions, with leading zero.

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

**Additional Improvements – “AI”** followed by 6 numeric positions, with leading zeros.

**Defensive Publication** – “T” followed by 7 numeric positions.

**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 leading zeros

**RX-Series Patents**– Reissued “X” Patents issued on or before June 13, 1848.

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

**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 7 numeric positions, without leading zeros. Patent Numbers with 6 or less positions will require leading zeroes. The 8th position of a utility fraction patent number will contain an alpha character as defined below

- Utility X Fraction Patent Numbers will contain “X” followed by 6 numeric positions, with leading zeros. The 8th 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 8th 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 8th 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 – ¼ L – 3/4

 E – 5/16 M – 13/16

 F – 3/8 N – 7/8

 G – 7/16 O – 15/16

 H – 1/2

 Example of Utility Patent and Fraction Patent Numbers:

02712152

2712152H

 02712153

**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*

There are one (1) type of Reexamination Certificate for kind code “B”:

(1). Ex Parte Reexamination Certificate under 35 U.S.C. 307

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

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

**Fn** – Supplemental Examination Certificate Published on or after September 16, 2012

## **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

**I*2*** *-* Reissued “X” Patents issued on or before 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 on or after September 16, 2012

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

**NP** - Non-Patent Literature

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

## **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