

MAPS™

Patent Text Data File Format Specification – Version 7.0c For Patent, Patent Application and Patent Abstract Text Document Formats

Published April 25th 2016 - IP Data Corporation

This also covers data element formats used in the MAPS-XML data files version 7.0 documented in DTD file named **IPData-maps-xml-v7.0-2015-09-01.dtd** available on the IP Data Corporation Web Site <http://www.IPDataCorp.com/>

Table of Contents

| | |
|--------------------------------------------------------------|-----------|
| INTRODUCTION..... | 6 |
| MAPS Versions and Changes in this Document..... | 6 |
| Changes in 7.0b (from 7.0a)..... | 6 |
| Changes in 7.0c (from 7.0b)..... | 6 |
| Purpose of this Specification | 7 |
| Multipurpose Data Storage Design..... | 7 |
| Acronyms, Abbreviations and Common Terms..... | 7 |
| Executive Summary of the MAPS Design..... | 9 |
| Ongoing Support for MAPS..... | 10 |
| Overlapping Support for Data Improvements..... | 10 |
| MAPS TECHNICAL DETAILS..... | 11 |
| Line Oriented Design..... | 11 |
| Paragraph Sizes..... | 11 |
| Character Encoding Schemes..... | 11 |
| Character formatting in MAPS (Bold, Superscripts, etc.)..... | 12 |
| ST.9 Standard Defining INID Codes..... | 12 |
| Table of INID Codes Condensed Data found in ST.9..... | 13 |
| References to ST.32..... | 15 |
| MAPS Codes..... | 15 |
| MAPS Codes Describe the Data..... | 15 |
| MAPS Codes with No Data..... | 16 |
| Some Older Codes Without Data Slowly Being Phased Out..... | 16 |
| Introductory Examples of Different Types of MAPS Codes..... | 16 |
| The Four Logical Sections of MAPS Documents..... | 19 |
| Outline Showing MAPS File Organization..... | 19 |
| MAPS File Contents by Kind Code..... | 20 |
| Data Types in Bibliographic Fields..... | 21 |
| Dates..... | 21 |
| Country Codes..... | 21 |
| Languages..... | 21 |
| Publication Numbers..... | 22 |
| Publication Numbers on Corrections Not Unique..... | 22 |
| US Publication Numbers | 22 |
| WO Publication Numbers | 22 |
| EP Publication Numbers | 22 |
| Publication Numbers in EP References..... | 22 |

| | |
|---------------------------------------------------------------|-----------|
| JP Publication Numbers | 23 |
| Conversion of Japanese Emperor Years to Gregorian Years..... | 24 |
| Defining Multiple Items on One Logical Line..... | 24 |
| MAPS Codes with Variations..... | 25 |
| MAPS Codes with Repeating Sets of Data | 25 |
| Description Text Section..... | 26 |
| Section Label Headings Logically Divide Description Text..... | 26 |
| Tables in the Description Section..... | 27 |
| Claims Text Section..... | 28 |
| Claims Statement..... | 28 |
| Dependent Claims..... | 28 |
| Tables in Claims Text..... | 29 |
| Search Report Text Section..... | 29 |
| Uniquely Identifying Publications or Inventions..... | 30 |
| Using the Publication Number in a DBS..... | 30 |
| The Application Number (or File Number) in a DBS..... | 30 |
| Variable Length Text Fields..... | 30 |
| Primary Languages..... | 31 |
| Languages in US Documents..... | 31 |
| Languages in JP Application Abstracts..... | 31 |
| Languages in EP Documents..... | 31 |
| Languages in WO Documents..... | 31 |
| DOCUMENT TYPES SUPPORTED IN MAPS VERSION 7..... | 33 |
| MAPS CODE TABLE..... | 34 |
| Columns In the MAPS Code Table..... | 34 |
| Element Types | 34 |
| Document Identification Group [10]..... | 35 |
| Publication Numbers..... | 35 |
| Publication Dates..... | 36 |
| Publication Kind Codes..... | 36 |
| Publication Language..... | 36 |
| Document Label (Name)..... | 37 |
| MAPS Version and File Creation Date..... | 37 |
| Document Identification Section MAPS Examples..... | 37 |
| Document Identification Section MAPS-XML Example..... | 37 |
| Correction Information [15] | 38 |
| Correction Information Examples..... | 38 |
| Application Information Group [20]..... | 40 |
| Application (File) Number..... | 40 |
| Application Date..... | 40 |
| Application Filing Language..... | 40 |
| Application Processing Language (EP only)..... | 40 |
| Family ID (from DOCDB)..... | 41 |
| Application Information Group Examples:..... | 41 |
| Priority Application Group [30]..... | 41 |
| Priority Application Data Entry..... | 42 |

| | |
|----------------------------------------------------------------------|-----------|
| Priority Application Group Examples..... | 42 |
| Public Availability Dates Group [40]..... | 43 |
| Technical Information Group [50]..... | 44 |
| Title Section..... | 44 |
| Classification Section | 45 |
| CPC Classifications..... | 45 |
| US Classifications..... | 46 |
| IPC Classifications..... | 46 |
| Classification Section Examples:..... | 47 |
| Patent and Non-Patent Cited References Section..... | 47 |
| US Patent and Non-Patent Cited Reference Examples..... | 48 |
| Related Documents Group [60]..... | 49 |
| Related Applications Section..... | 49 |
| Concerned Parties Group [70]..... | 52 |
| Opponents and Our IID Database..... | 52 |
| Applicant Section..... | 53 |
| Inventor Section | 54 |
| Assignee Section..... | 54 |
| Agent Section..... | 54 |
| Examiner Section..... | 55 |
| Examiner Section..... | 56 |
| PCT and National Phase Group [80]..... | 56 |
| Designation States Section..... | 56 |
| International Filing Information Section | 58 |
| Patent Cooperation Treaty (PCT) Applications..... | 58 |
| Hague Agreement..... | 58 |
| International Number Formats..... | 58 |
| PCT Publication and Application Number Section..... | 59 |
| Search Report Date Section..... | 59 |
| Abstract Text Group..... | 61 |
| Description Group..... | 61 |
| Description Section and General Text..... | 62 |
| Claims Text Group..... | 63 |
| Claims Section..... | 63 |
| Search Report Group..... | 66 |
| Search Report Section | 66 |
| Family Document Section..... | 66 |
| General Text Codes Used in One or More Sections..... | 67 |
| End of Patent..... | 68 |
| Miscellaneous Comment Entries (CMNT)..... | 69 |
| End of MAPS CODE Table..... | 69 |
| APPENDIX A – NAME AND ADDRESS DATA FOR CONCERNED PARTIES..... | 69 |
| Background Information..... | 69 |
| MAPS Party Fields with Name and Address Data..... | 70 |
| How the Party Fields are Divided..... | 70 |
| Name (Person or Entity/Organization)..... | 70 |

| | |
|-------------------------------------------------------------------------------|------------|
| Address..... | 70 |
| City, State and Region..... | 70 |
| Country per ST.3 Standard..... | 70 |
| Postal-Code..... | 71 |
| Additional Elements in (APLX) Applicant (all authorities)..... | 71 |
| Additional Elements in EP (INVX) Inventor..... | 71 |
| Additional Elements in U.S. (ASGX) Assignee..... | 72 |
| Additional Elements in (APLX, AGTX and ASGX) Applicant, Agent, Assignee | 72 |
| APPENDIX B – COUNTRY AND LANGUAGE CODES..... | 74 |
| Country and Language Codes listed in MAPS patent documents..... | 74 |
| Country Codes | 74 |
| Language Codes..... | 74 |
| ISO 639 Codes (Names of Languages)..... | 74 |
| APPENDIX C – CORRECTION CODES FOR PCT (WO) PUBS..... | 75 |
| PCT Correction Codes used in ALL MAPS WO Pubs..... | 75 |
| Table of Correction Types (single digit, 1 to 4)..... | 75 |
| Table of Republication Codes (R1 to R9)..... | 76 |
| APPENDIX D – DOT CODES IN OLDER US DATA..... | 77 |
| Introduction..... | 77 |
| Background..... | 77 |
| Subset of Green-Book Appendix | 78 |
| USING THE DOT CODES..... | 79 |
| On the USPTO Web Site..... | 80 |
| The DOT CODE SYMBOL/CHARACTER TABLE..... | 80 |
| APPENDIX EP – EPO SPECIFIC INTERNAL TAGS | 91 |
| APPENDIX H..... | 93 |
| USPTO, EPO and WIPO Patent-Data Related History..... | 93 |
| THE USPTO, APS and the 1970s and 80s..... | 93 |
| The late 70s and the EPO and WIPO Get Involved..... | 94 |
| APPENDIX K – KIND CODES, ALL AUTHORITIES..... | 96 |
| Kind Codes for PCT and EP Documents..... | 96 |
| Kind Codes for US Patent Documents..... | 97 |
| APPENDIX L – EP AND WO LANGUAGES IN MAPS..... | 99 |
| EP Application Languages..... | 99 |
| EP Granted Patent Languages | 99 |
| WO (PCT) Application Languages..... | 100 |
| APPENDIX M – MATH-ML DATA | 101 |
| Typical Browsers and Math-ML..... | 101 |
| Browsers That Can't Display Math-ML Without Modifications..... | 101 |
| Browsers That CAN Display Math-ML Without Modifications..... | 101 |
| Math-ML Files included as Separate HTML-5 files..... | 101 |

| | |
|-------------------------------------------------------------------------|------------|
| Using a Generic Inline frame to display Formula Files..... | 102 |
| Why We Pulled the MathML From the Files..... | 102 |
| Building the Math-ML Filename..... | 102 |
| MathML with Default Math Fonts on Firefox, Safari and Opera..... | 103 |
| A Bug in Mozilla and Opera Browsers with the Default MathML Font..... | 103 |
| Safari Browsers (Mac and PC both) look OK Except for One Character..... | 103 |
| Some MathML Uses Newer, Not-yet-supported Character Codes..... | 103 |
| Substituting the Older for the Newer Characters, If Needed..... | 105 |
| StringSubMany Utility for Mass Text Substitution (Windows)..... | 105 |
| APPENDIX P - FOR THE PROGRAMMER..... | 106 |
| Useful Tables, Data Sets and Code Samples..... | 106 |
| Utility Programs to Manipulate & Verify MAPS Data..... | 106 |
| APPENDIX S - SEARCH REPORT TEXT DATA FORMAT..... | 107 |
| Two Main Sections in the International Search Reports..... | 107 |
| Example of Search Report Text | 108 |
| APPENDIX T – ST.8 STANDARD CLASSIFICATIONS..... | 110 |
| Details on USPC, CPC and IPC Class formats..... | 110 |
| CPC and IPC Both Use ST.8 Standard..... | 110 |
| Only 15 of the 50 Bytes Define the Technology Class..... | 110 |
| Display and Storage of the 15 Class Characters..... | 110 |
| Class Data Format Table from ST.8 WIPO Standard..... | 111 |
| APPENDIX U – UNICODE UTF-8 CHARACTER SETS FOR MAPS..... | 113 |
| Supported UTF-8 Ranges (all Languages)..... | 113 |
| Indexing UTF-8 Data with Symbols and | 114 |
| APPENDIX W – WIPO PUBLICATION NUMBER DETAILS..... | 116 |
| Introduction..... | 116 |
| PCT Publication Number Formats from 1978 to Now..... | 116 |
| Initial PCT Publications Number Format in 1978..... | 116 |
| PCT Publication Sequence Number Goes to 5 Digits in 2002..... | 117 |
| PCT Publication Number “Modernized” in 2004..... | 117 |
| EPO Accepts Some PCT Applications as Native EP Applications..... | 117 |
| EuroPCT Applications..... | 117 |
| Prior to 2009 WIPO Does Not Use Correction Kind Codes..... | 118 |
| EPO Renames WO (PCT) Correction Documents (new Kind Codes)..... | 119 |
| Additional Details on PCT (WO) Kind Codes..... | 119 |
| PCT (WO) Primary Documents..... | 119 |
| Correction Documents..... | 119 |
| PCT Amended Claims Use Kind Code A4 (or B1 prior to 2009)..... | 119 |
| APPENDIX Z - ABBREVIATED TABLE OF MAPS CODES..... | 121 |

INTRODUCTION

MAPS Versions and Changes in this Document

The MAPS Specification is now in its 7th major version since its inception in 2003.

MAPS version numbers are composed of a **major version number** which is one or more digits, a **minor version number** after a decimal point, and end with a **minor revision code**, which is a lower case letter.

The current version is **7.0b** which is a minor documentation update for Version **7.0a** which replaced Version 6.3f in its entirety. The background color used for this paragraph is used on all changes found in the **7.0b** specification. **The MAPS-XML DTD was not affected** and remains at Version 7.0 dated 2015-09-01.

The current version is **7.0c** which is a minor documentation update to Version **7.0b**. The background color used for this paragraph (yellow) is used on all changes for the **7.0c** specification. **The MAPS-XML DTD was not affected** and remains at Version 7.0 dated 2015-09-01.

An increase in the Major version number indicates that all Minor version changes and Minor revisions are being incorporated in the manual in the proper sections with all change indicators removed. Plus, a change in the Major Version number can indicate that some data element has been added or modified such as it may behoove data users to update their inventories with copies of the data in the new version.

Active subscribers for 2 or more years, or those who purchased the backfile, will receive a complete copy of the new data for the Authorities to which they subscribe.

Future changes in this document will be highlighted. Previous *change indicators* using left margin change bars are no longer used.

Changes in 7.0b (from 7.0a)

The minor revision code had changed from (a) to (b). A correction to the specification has been made which involves MAPS code **REIS** in the **Related Documents Group** of the MAPS Code Table. This code was not properly documented. The **REIS** code is only found in Reissue Patents published prior to 2001. The MAPS Code Table has been corrected and additional examples have been added to support the change.

Version 7.0a of the MAPS specification was not yet provided to all users (Version 7.0b is the for full distribution).

Changes in 7.0c (from 7.0b)

The minor revision code had changed from (b) to (c). A correction has been made to document additional information on four MAPS codes in the PCT Publication Section (INID Group 80). Early in 2015, changes were made to US law for the **Hague Agreement** which allows filing a standardized “international” application for Design patents. Like the PCT, WIPO accepts and manages applications and maintains a database for them which also lists designated contracting countries. A U.S. Application must be filed in parallel to the International application and is assigned a U.S. Application number using Series Code (35). Like the PCT data, the USPTO includes 4 pieces of information (numbers and dates) with each Design patent that was filed under the Hague Agreement. **We store this new data in existing MAPS codes that serve the same purpose under the PCT.** The affected codes are **IPNO**, **PPDT**, **PCFD**, and **P102**. The details are described in the MAPS Code Table beginning on page 58. In Appendix S, figures was not capitalized and missing a colon (e.g., “figures” was supposed to be “Figures:”). The background color should make them easy to find. A few minor “typos” elsewhere in the specification were also corrected, plus two date examples were added which have no affect on content, but are also highlighted.

Purpose of this Specification

This specification defines a text-based data storage and transfer format named MAPS™ used for granted patent and patent application text, and various update documents. The specification covers full text versions as well as bibliographic versions with abstracts (all referred to in this specification as **the documents or the publications**).

This defines the file storage format, character encoding requirements, supported data elements (fields), and supported languages. Some text data are included in languages that we technically do not support, and this data are passed on “*as received character for character*” from the various patent authorities. The single exception that the character set may be converted to Unicode UTF-8, which is now **the only set we support**.

Multipurpose Data Storage Design

The **MAPS** format is a *text-based, line-oriented* data storage format that defines a set four-character identification tags called **MAPS Codes**. The codes identify line oriented-data fields in the MAPS documents.

The **MAPS** design provides:

- 225 years of patent data from multiple authorities in the same easily-parsed format.
- Easy corrections with common UTF-8 text editing tools (required to support OCR data).
- General purpose patent data exchange or static storage.
- Easy electronic distribution.
- Easily defined database parsing and indexing.
- Reduced file sizes
- Line oriented parsing for minimal programming for DBS indexing and accurate display on a very wide variety of systems.

Acronyms, Abbreviations and Common Terms

The following acronyms, abbreviations and phrases are used throughout this document. Some of them may be discussed in greater detail in later sections of this specification:

Abstract – A concise one or two paragraph description of an invention.

Bibliography – (a.k.a. “the Bib” or “Front Page Data”) - A collection of data elements usually presented on the front page of complete granted patents or patent applications. Also found on further publications for updates or changes to the legal status, and on abstracts of the originals.

The elements usually include:

document identification (publication number, date published),
application or filing information (Authority assigned File number and date),
priority application lists,
classification data,
parties involved in the application (applicant, inventors, agent, assignee, etc.),
contracting states (PCT and EP publications),
patents and non-patent literature (referenced in the patent), and others.

Applications and patents are often available in an abbreviated format which includes the bibliographic data and an abstract text paragraph, and are usually called an “Abstract.”

CPC – Cooperative Patent Classification (system) – A new classification system shared by agreement between the USPTO and the EPO to standardize on a common set classifications with the format based on the ST.8 WIPO Standard. Prior to the CPC, the EPO used ECLA, discontinued as of Jan 1st 2013, and the US used the US Class system, ending Dec. 31st 2015. A small part of the US system is still used for Design patents; which are not included in the CPC.

EPO - European Patent Office - A multinational agency located within the European Union (EU), in several locations, responsible for administering the **European Patent Convention (EPC)**. A patent system created by treaty among EU nations that provides patent protection in one or more of the Contracting States (designated individually). Additional functions assumed by the EPO include maintenance of a database of almost all patents from all nations around the world (eSpace system) along with a continually updated database of the Bibliographic data and most abstracts of these same patents (**DOCDB**). Plus, they also maintain the **INPADOC** database (**a.k.a. PRS**). INPADOC tracks legal status of the world's patents as a series of legal transactions, as defined and updated through weekly submissions by each participating nation. This is one busy group of approximately 7000 employees (as last count).

Further Publication – A document that is published *in addition to* the Primary or First Publication documenting an invention; which is usually the first Published Application. Further publications document legal changes, a correction, a completed Search Report, or additional work that has been completed prosecuting the initial application up to and including a granted patents (if there is one). Most countries assign Kind Codes to assist with the identification of these Further Publications. For countries that use the same sequence number for all publications dealing with one invention, the Kind Code is not just important, it is mandatory on Further Publications.

IPC – International Patent Classification (system) – A patent classification system maintained by **WIPO** (see later definition). Its format is also based on the ST.8 Standard. The IPC system is used by most patent authorities as a secondary class system on their publications in addition to their National system. In many cases the IPC class may be the only classification initially found on a publication because detailed classification in more complex systems may take more time. An IPC classification may also be required under the PCT (see later definition of PCT).

ISR – International Search Report – A report generated as part of the PCT process usually included with Kind **A1** PCT published applications in printed form, or when delayed past the first required publication date for an application. The Kind A1 becomes a Kind A2, and the delayed Search Report becomes a *Further Publication* as a Kind A3 document (in other words, A2 + A3 = A1). Even though the Search Report is included in the printed versions of A1 and A3 Kind documents (and therefore the scanned image copies of these), WIPO and the EPO are experimenting with providing search report data with the electronic data they distribute. When available, they are currently from OCR. Although, the EPO indicates they are working towards an electronic source, and we assume (and hope) that WIPO is also working to that end.

MAPS is an acronym for **Modified APS**. “**APS**” was the US Patent & Trademark Office's venerable **Automated Patent System**, designed circa 1970, and used for over 30 years (very few computer software systems are in use that long). For further information, see **Appendix U, USPTO Patent Data History** in this document.

Patent Documents – an all-encompassing term referring to published patent applications, granted patents (“the grant”), corrections, amendments, and search reports.

PCT – Patent Cooperation Treaty. An international treaty administered by WIPO (see next definition) ratified by more than 140 countries, the goal of which is to provide a standardized patent application acceptable to all signatories. Under the treaty, after an international application is successfully filed, the applicant may designate or elect any of the signatory countries to seek patent protection at the national level. Further prosecution and the granting of patent protection remains the responsibility of each national patent office.

WIPO – World Intellectual Property Office. The United Nations agency responsible for administering the PCT. The agency, also known as the **International Bureau (IB)**, its abbreviation, is also its designated Country Code in ST.3) is responsible for the following:

- receipt of record copy (official applications, corrections, etc.),
- receipt of amendments to the claims,
- International publication of the applications, corrections, and amended claims,
- distribution of the applications and ISRs to designated Offices,
- communication of the Preliminary Report on Patentability,
- maintain a series of published standards, Applicant's Guide, a PCT Glossary, and
- maintain a database of all PCT applications and related paperwork under the Treaty (known as **PatentScope**).

Executive Summary of the MAPS Design

MAPS is a simple, technically accurate, line-oriented text data format designed to store and convey patent document text currently covering data collections from the US, EU, WIPO and Japan (English Abstracts of unexamined applications). The collections are all converted to UTF-8 characters for all languages, including scientific symbols, characters, math and logical operators. It is easily edited which is required for us to support OCR supplied by some Authorities, and to support OCR created by us for repairs or additions to older inventories. The line oriented text facilitates easy parsing for indexing, and for display in a wide variety of text-oriented database applications, and browser-based Internet based applications.

Any text indexing system that allows you to define beginning and ending points to delineate document sections (fields, elements or text areas) can easily be set up to index MAPS files. This especially includes the newer free-form text indexing and search systems which allow people that know about patents, not necessarily how to program computers, to create very powerful databases with MAPS data.

Below are examples of some common MAPS fields. The MAPS Codes, shown in **BOLD** typeface in the following examples, begin each line followed by the data for that code. Most MAPS codes that indicate section boundaries do not have data:

| Code | Data | (Comment – not part of MAPS data) |
|-------------|-------------------------------------------------------|--------------------------------------------------------|
| PATB | US9100125B2 | (PATB indicates a US Patent w/ the publication number) |
| PBDT | 20150804 | (PBDT is the Publication Date) |
| APLX | ViaSat Inc. 6155 El Camino Real Carlsbad, CA US 92009 | (an Applicant) |
| INVX | Fan, Mo Hinkley, OH US | (an Inventor) |

Providing all data using Unicode UTF-8 character encoding allows us to support over 30 languages. IP Data also produces a **MAPS-oriented XML** version of the complete collections from each of the Patent Authorities we support that we generate *on demand*. This is Valid XML and conforms to our publicly available DTD specification with version numbers and changes that parallel the MAPS versions.

Ongoing Support for MAPS

MAPS has been in existence for over 13 years. When an issue arises with the data, we attempt to solve it immediately – *even if it requires us to replace 10 million files (we have done this several times – even for cosmetic repairs)*. Sometimes we do it right away, sometimes not. Delays are usually due to errors that are beyond our control (e.g., corrupted or missing patent authority data).

If additional engineering is required to help you solve a problem on your end with your system that handles the MAPS data; we are always happy to see if we have already solved it in one of our systems. We can see if another MAPS user has an answer (*and we will be happy to share it, with permission, and as long as it does not violate someone's intellectual property rights*).

Contact IP Data Corporation for more information on utilities available to support large MAPS inventories. We also have coding examples (in the C programming language) to parse UTF-8 data and display MAPS files using HTML.

Overlapping Support for Data Improvements

When MAPS data fields change where the amount of data has increased, such as added detail from the patent authorities, rather than change a long understood MAPS line format, we may add a new MAPS code with the updated data and keep BOTH fields in the distributed data, if it make sense to do so. We have also opted to create two versions of the documents and publish both old and new, side-by-side, to give users time to make updates and changes in their systems. As long as data is available to support an older format, we can do this as long as necessary to support your needs (we have published parallel formats for as long as 14 months in the past). The only conditions where we can't do this is when the older data field is in error or has been removed. We will not knowingly distribute incorrect or misleading data. We will notify subscribers if this happens as soon as we know about it.

MAPS TECHNICAL DETAILS

Line Oriented Design

MAPS is a line-oriented design. Which means that each MAPS file is a collection of lines made of 8 bit bytes, each ending with an End-of-Line (EOL) character sequence. The default sequence is two bytes known as Carriage Return and Line Feed (CR, LF) with hexadecimal values 0Dh 0Ah (13 and 10 decimal), and a single line feed is also acceptable as an EOL Sequence for MAPS files. MAPS-XML uses a single Line Feed for the EOL (0A Hex). This is the traditional EOL used in the Linux and Unix[®] worlds, and is the recommended EOL in XML. MAPS parsers should handle both EOL situations.

MAPS Codes are always four single byte ASCII characters, and always begin with an UPPER CASE letter. While the remaining three characters may be any combination of capital letters and digits. The code is followed by a space (ASCII 20 hex), **then optionally by data** as defined in the specification for each of the different codes, and the line always ends with the EOL sequence.

If you open a MAPS file in a text editor with Word Wrap turned off; you will see that every line begins with a MAPS code. The data for each line extends to the right until an EOL character sequence is reached.

Paragraph Sizes

Some Authorities ship patent data with **extremely long paragraphs**. Many over 100 kilobytes, quite a few as long 275 kilobytes, and a few we have handled that exceed one megabyte. These are not tables, but are confirmed to be text paragraphs with things like lists of diseases, lists of chemical strings, lists of values for numeric arrays and matrices and names of various items. **We hand edit these unformatted lists** in our “modified *source copies*” of the individual source XML or SGML text. We break them into more manageable sizes making them groups of paragraphs, all approximately the same size, and usually from 5000 to 6000 characters each. Not all long paragraphs are divided into smaller paragraphs, but most that are found to be longer than of 16,384 characters are at least cut in half into the two paragraphs. **Please note that 16,384 characters does not mean 16,384 bytes since some UTF-8 characters are composed of 2, 3 and occasionally 4 bytes for a single visible character. Chinese, Japanese and Korean (CJK) Ideographic characters are all 3 bytes long encoded as UTF-8.** This means that parsers should be prepared to handle large paragraphs; some as large as 40 to 50- kilobytes.

Character Encoding Schemes

Most database platforms and some programming languages (through their support libraries) handle multi-byte characters “*behind the scenes*” allowing the programmer to deal with Unicode “*Code Point*” values (32 bit integers) written as U+00001 to U+nnnnn with nnnnn representing the highest supported code. The values appearing in technical documents usually use hexadecimal values (e.g., Unicode+010AB).

The Unicode values are numbers assigned to the characters that never change. This allows multiple character encoding schemes to be used without changing any of the programmer's code or adding any additional support features to the database product.

When working with MAPS or MAPS-XML data, always specify UTF-8 for your encoding configuration.

We chose to convert directly to UTF-8 binary data eliminating all known Ampersand-Entities in the data for three reasons:

1. It is the most efficient character encoding design for multiple languages, **bar none**.

2. XML uses UTF-8 as its preferred, default character encoding scheme.
3. In the early implementations of SGML and XML used for patent data by the various Authorities, named HTML Ampersand Escape sequences were used everywhere. It was not uncommon to find two or three different named *escape sequences* for the same Latin diacritical character, and more often than anyone will admit, some, or all of them were wrong. The sooner we got copies of the original data into UTF-8 (for easy editing) and displayed the “legal copy” (TIF images, in our case) to correct all errors including errant escape sequence, the fewer times we had to deal with it if we had to regenerate sections of our inventories from the original source data.

Character formatting in MAPS (Bold, Superscripts, etc.)

Older US and PCT data may still contain APS “Dot Codes” for character formatting, superscripts, subscripts and for some special characters and symbols, although less than in prior MAPS versions since we are actively replacing them with UTF-8 characters and the subset of HTML we use for simple highlighting. It is taking far longer than we anticipated since all exceptions to our automated process must be viewed and compared to the image copy (Legal copy). For information on the APS Dot Codes, see **Appendix D**, titled **Dot Codes in Older US Data**.

Newer MAPS files (2001 and later) contain HTML for basic highlighting and for most Superscript and Subscript characters. We will be converting many of the all-numeric super and subscript characters into UTF-8 character codes that directly display the characters without HTML tags. Instead of seeing CO² in your editor, it will see numeric superscripts and subscripts as you would any other characters, like this: CO². The superscript 2 is Unicode character (U+00B2, **C2 B2** UTF-8 binary) and we will begin to use the numeric values wherever we can. The text is much easier to read and edit using these codes for the superscript and subscript digits. It allows us to more easily correct chemical strings and other complex math data.

The only HTML character formatting used is:

- Italics <i>Italics</i>
- Bold Bold
- underline <u>Underline</u> (no attributes, single underline)
- subscript Sub_{_{script}} ==> Sub_{script}
- superscript Super^{^{script}} ==> Super^{script}
- strike-through gone (is the HTML 5 version of <strike>, already supported in all browsers)

The simple HTML tables that were used in the last year or two of data in some of the collections **have been completely eliminated** in favor of accurately spaced, line-oriented plain text where spaces are significant for columnar data alignment. This also means that **<pre> Table Data </pre>** should be used for display in HTML server systems and is supported in all browsers.

ST.9 Standard Defining INID Codes

INID is an acronym for: “**Internationally agreed Numbers for the Identification of (bibliographic) Data.**”

The INID data is published by WIPO as **International Standard ST.9**. The INID system was originally designed for use on the face of patent documents to help readers (human, not computers) to identify each of the data elements they see. You will see these codes in a smaller font inside square brackets on the face of almost all patent documents around the world.

Since it “neatly” lists and organizes all of the bibliographic elements (i.e., front page data) grouped numerically, we make use of it organizing and aiding in the identification of the data elements contained in

MAPS files. Merging the data from multiple authorities is no small chore, and organizing our data around the INID numeric structure is an excellent way to do this.

We make use of it in two ways:

1. We organize the data in the order provided in the INID Standard (with some exceptions), and
2. We list the INID code with each defined MAPS code in the main MAPS code table, later in this specification.

We have created the table below condensing the information from the WIPO Standard for our use, and for your convenience. We use these numbers in our **Outline Showing MAPS file Organization**, below, and again for each element in our **MAPS Code Table**, also below, (as applicable) to assist with our definitions. Please check the WIPO Standard for updates when making important decisions concerning **INID Codes**. Like you, we are only another user of the data and have no authority concerning the actual definitions or additions to the standard. But, we do make it our business to help you, our users, understand exactly what we deliver to you.

Table of INID Codes Condensed Data found in ST.9

* Minimum Required on the Front Page of a Printed Patent Publication (according to the standard).

| INID No. | Element Title | Description (if needed) |
|--------------|--------------------------------------|------------------------------------------------------------------------|
| 10-19 | Publication ID Group | Element Group Identifying the publication (Patent, Appl. Etc.) |
| 11* | Publication Number (national) | (10) may be used if 11, 13 and optionally 19 used on one line |
| 12* | Publication Type (text name) | e.g. "US Patent" - "International Application under the PCT" - etc. |
| 13* | Kind of Document (a Code) | Kind Code from WIPO ST.16 - (A, A1, A2, B2, etc) |
| 14 | Date Published | Date of Publication of document in/on which it resides. |
| 15 | Correction Data | Data identifying document as a correction (as a minimum). |
| 19 | Organization | Who published the document (Country Code from ST.3) |
| 20-29 | Application Data Group | (21-29) Data Concerning the Application |
| 21* | File (or Application) Number | Original File Number Authority Assigns upon application |
| 22* | File (or Application) Date | Date Application was Filed with the Authority |
| 23* | Other Important Date(s) | Publication Date of Spec - date of full filing after provisional, etc. |
| 25 | Filing Language (NOT country code) | 2 lower case letters from ISO-639 (Int'l Language abbreviations) |
| 26 | Publication Language | Primary Language of the document (ISO-639) |
| 27 | Mention of Previous Application | To secure File Date under PCT Art. 5(7) |
| 30-39 | Priority Data Group | (30-39) References to Publications having Priority |
| 31* | Priority Appl. Numbers | (30) maybe used if all items (31 32 33) are in one item |
| 32* | Priority Appl. Dates | Date of the application listed in item (31) |
| 33* | Country of Pri. Application | (2 Letter state, in ST.3) - "WO" may be used for Int'l Apps |
| 34 | Regional or Int'l Pri. Agreements | One or more countries who are parties to Paris Convention |
| 40-49 | Public Availability Dates Grp | (41-49) Dates when Appls. made public in various phases |
| 41 | Unexamined Patent w/o grant | Unexamined Patent Published without Grant |

| | | |
|--------------|------------------------------------|----------------------------------------------------------------------|
| 42 | Examined Patent w/o grant | Examined Patent Published without Grant |
| 43 | International Publication Date | Unexamined Application Published (should be 1st public access) |
| 44 | Examined Patent w Provisional | Examined Patent Published with Provisional Grant only |
| 45 | Granted Patent Published | Patent published with Grant on this date |
| 46 | Patent Claims only Published | Date on which only Claims of a Patent were Published |
| 47 | Date Patent avail. on Request | Patent with Grant made available upon request this date |
| 48 | Pub Date of Correction | Date a Corrected Version of the document was published. |
| 50-59 | Technical Information Group | (51-59) Various fields assigned to describe or classify Appl. |
| 51 | International Class | (IPC) International Patent Classification |
| 52 | National Class | Class assigned by Country of patent (e.g., CPC for US or EP) |
| 54 | Title | Title text of Invention (patent or application) |
| 56 | References Cited | Both Patent and Non-Patent related documents Cited |
| 57 | Abstract Text | Paragraph (MAPS locates this last in the Bib Section) |
| 58 | Field of Search | Technology Field(s) searched, listed broadly by Class |
| 60-69 | Related Domestic Doc Group | (61-69) Related Applications and Patent (from same Authority) |
| 61 | Related Prior Appl. Or Doc. | To which the present document is an addition |
| 62 | Divisional Related Appl. | Earlier Appl. from which present appl. is divided from |
| 63 | Continuation Related Appl. | Earlier Appl. from which present appl. is a Continuation of |
| 64 | Related Reissue | Earlier Appl. which is Reissued (may also contain pub. number) |
| 65 | Prior Publication Data | Publ. Number of earlier version of this doc (e.g., A1 for this B2) |
| 70-79 | Concerned Parties Group | Identification of parties associated with this document |
| 71 | Applicant (Except US before 2014) | Applicant – where laws do NOT require Applicant be the Inventor |
| 72 | Inventor | For US use INID 75 (required under US law) |
| 73 | Assignee - Grantee | For US use (76) – Grantee Assignee(s) Holder(s) or Owner(s) |
| 74 | Agent – Legal Representative | Register Agent - Attorney or Law Firm representing applicant |
| 75 | Inventor/Applicant | U.S. Only when laws required Inventor be applicant (prior to 2014) |
| 76 | Inventor/Applicant/Grantee | U.S. Only, when Inventor was also Initial Assignee (prior to 2014) |
| 80-89 | PCT and National Phase Info | PCT Application and Document Group |
| 81 | Designated States | States designated for Patent protection (WIPO/EPO) |
| 84 | Designated Regional States | Regional States designated for Patent protection (WIPO) |
| 85 | Date Begin National Phase | Not usually included in Data supplied by Authorities |
| 86 | PCT Filing Data | Application Number, Date Filed with primary authority |
| 87 | PCT Publication Data | Publication Number and Publication Date (by WIPO) |
| 88 | Search Report Date | Date International Search Report was completed |

You may acquire ST.9 along with other recommended standards from the World Intellectual Property Office's web site. As of the this date, the list of standards may be found at this Internet address:

http://www.wipo.int/standards/en/part_03_standards.html (link validated January 2016).

References to ST.32

You may see references to the ST.32 standard where tag names are concerned. This standard defined fields using an alphanumeric code that appears to be based on an expanded version of the INID codes. The tags in the current EP files (ST.36) use most of the older ST.32 codes. The US also used them in very early SGML and XML data (2001 to 2004 granted patents), but moved away from it in favor of long, descriptive word-based tag names with the only drawback being the increased file sizes.

For the benefit of those familiar with ST.32 tags, we list the ST.32 code, when known, below the INID number in that column of the MAPS code table, later in this guide. It is not a problem if you are not familiar with it. We try to add as much information for users with all levels of experience dealing with patent data. Examples of the ST.32 tags: **B100** - Publication ID Group, **B200** – Application ID Group, **B700** – Parties Group, **B710** – Applicant, **B720** – Inventor, etc.

MAPS Codes

Maps codes are all four characters long, and always begin with a Capital letter (A-Z). The last three characters must be a capital letter (A-Z) or a digit (0 to 9). These characters must be the first four characters on a physical line. No blank lines are allowed except after the last MAPS code which must be **PATE** (End of MAPS data) which ends the file parsing.

When parsing a MAPS line, the first thing to check is the validity of the MAPS code, itself. Described in a simple, pseudo-programming language, this is the logic to validate a MAPS code:

mapscore[1] to **mapscore[4]** represent the four characters in a MAPS code.

```
IF ((mapscore[1] is A-Z) AND
    ((mapscore[2] is A-Z) OR (mapscore[2] is 0-9)) AND
    ((mapscore[3] is A-Z) OR (mapscore[3] is 0-9)) AND
    ((mapscore[4] is A-Z) OR (mapscore[4] is 0-9)) AND
    (character after maps-code is-a-space))
```

THEN

The MAPS code is valid, and data may follow based on the specific code

ELSE

The MAPS file is corrupted (or is not a MAPS file)

MAPS Codes Describe the Data

Each MAPS code is followed by a space, and optionally by data. The MAPS codes determine exactly what type of the data follows them on the same physical line. **MAPS-XML** uses the same codes for data, or a group element tag name that contains more than one element if the MAPS code had data with multiple elements such as a Publication Number and a Date. The XML tag names are lower case versions, as is customary, though not mandatory for XML tags.

MAPS Codes with No Data

MAPS codes with no data on the same line are usually Section Dividers except for some **TABL** lines that patent text authors add for the visual appearance in a table. But, not all Section Dividers are without data. For example, the **CLMS** code (Begin Claims Section) has an optional Claims Statement such as “*We Claim:*” or “*What is claimed is:*” (optional because some patent authors don't add a claims statement).

Some Older Codes Without Data Slowly Being Phased Out

Earlier versions of MAPS defined several section-begin codes that also indicated the existence of some heading text on a face of a printed patent. These are no longer needed, and we are slowly phasing these out during 2016. Some of our files may still have them, but they will disappear over the 2016 calendar year as we update various code modules that generate them.

The following tags are still defined, but will slowly disappear over 2016:

| MAPS Code | Heading text | Original Purpose |
|-------------|----------------------|----------------------------------------------------|
| REFS | “Cited References” | Begin Cited References (if any existed) |
| URFS | “U.S. References” | Begin U.S. patent references in a US patent |
| FRFS | “Foreign References” | Begin Non U.S. patent references in a US patent |
| ORFS | “Other References” | Begin Non-Patent References in a US and EP Patents |

Earlier US patent data (Green Book format) did NOT have the “US” Prefix on the patent numbers in the references, and the **URFS** Section-Begin code was required to indicate that the following numbers were all US patent references. It has been several years (and several MAPS versions) since US patent references were not prefixed with “US” and non-U.S. patents have always had the country prefix on them. The **ORFS** code will also be phased out since all NPL (Non-Patent Literature) References use the unique **OREF** code and are always grouped together. They appear directly after Patent References end, if any existed.

Since all patent references are grouped together, your display logic can emit any heading text it desires when you read the first patent reference, and another heading for the NPL references when the first one is read.

Another change that will happen later in 2016 is the **UREF** and **FREF** codes will be discontinued from U.S. Patents and all patent references in all documents will begin to use the MAPS Code **REFP** (which is the preferred single **Cited Patent Reference** MAPS code).

Introductory Examples of Different Types of MAPS Codes

This section is only an introduction to MAPS codes. The **Maps Code Table** that begins on **page 34** defines all MAPS codes in detail with examples of use, and you will also find a more detailed discussion on the MAPS text paragraph code types and claims text codes discussed in the sections titled **Description Text Section** and **Claims Text Section** that begin on page 26 of this specification.

A few examples of MAPS codes, with and without data, are shown below after a description of each one. The MAPS line is shown in a fixed width font with the MAPS code in **bold** and the end-of-line character

code sequence shown as `<eol>`):

All MAPS files begin with the Publication Number code. **It is always the very first line in a MAPS file.** These are codes such as **PATA** for US Applications, **PATB** for US Patents, **PATU** for all EP documents, **PATW** for PCT World Applications and **PAJA** for Japanese Abstracts (in English). Even though these codes only have the publication number as data, the MAPS code indicates the type of document (i.e., what sections it will contain), and the first two letters of the Publication Number also indicate the patent authority that published it. Here are some examples:

PATA US20010000001A1`<eol>`

PATB US8137811B1`<eol>`

PATW WO2015001847A1`<eol>`

All publications with **Full Text** contain the **Description** section which begins with code **DESC** that has no data, and is on a line by itself. We add a single space character after all MAPS codes whether or not they have data. In the case of MAPS codes with no data, the single space will exist in the files before the End-of-line sequence. Our MAPS validation software looks for this space, but we do know of text editing software that remove trailing spaces on lines when they save the file. For this reason, the space is not mandatory for any purpose other than our internal validation software and when data follows it on the line. There will ALWAYS be a space after the MAPS code before any data begins. The **DESC** code implicitly ends the Bibliographic section of the document:

DESC `<eol>`

DTXT is a text paragraph in the *Description* Section of a Full Text MAPS file. The text in the example as shown is “wrapped” by the word processor software (creating this spec.) and contains no control characters before the End-of-Line character sequence shown as `<eol>`:

DTXT The preferred embodiment is only one example of use and should not be construed to be the only embodiment which would limit possible applications not described here in detail. `<eol>`

Full Text Publications also contain a Claims Text Section. Claims Text sections usually begin with code **CLMS**, except for EP documents where the language of the Claims is indicated by one of the three codes with **CLMS** for English, **CLMD** for Deutsche or German, and **CLMF** (for French). EP Applications will have one of the three codes depending on the language, while EP Granted Patents are unique in that they have Claims in all three languages in the same document. How multiple languages are handled in EP and WO documents is discussed in details in **Appendix-L** titled **EP and WO Languages in MAPS**.

The Begin Claims codes also have Optional Data which is the *Claims Statement* by the Inventor or Applicant, and is usually a short statement like “**I claim:**” or “**We claim:**” or “**What is claimed is:**”, for example:

CLMS I Claim:

CNTX is a numbered claim element in a patent or application, while **CTXT** is an unnumbered Claims Text Element and used for claims where the application author has divided a claim into two or more sections, for example:

CNTX 1. A system and method to store and process patent related text, including: <eol>
CTXT a series of predefined four character codes, one for each type of data; <eol>
CTXT a single, standardized character encoding scheme shared by all data; and <eol>
CTXT a series of software programs to process 21 different formats converting them into a single, easy to use format allowing both computer programmers and non-computer programmers to make more efficient use of the most popular patent data from around the world. <eol>

PBDT is the publication date of the document, while **ISSD** is the also a publication date, but also the date a grant of protection is **ISSUED** for a patent (all dates are formatted: YYYYMMDD):

PBDT 20130107 <eol>
ISSD 20130107 <eol>

The **ISSD** code is a “holdover” from the Green Book data, and **PBDT** will also work in a patent to indicate the Publication Date for a Patent as well as a Patent Application. We think that having a patent granted is something very special, and this is why we have been reluctant to simply eliminate the **ISSD** code.

PATE is the End Patent Data Code and all MAPS files end with this code. This ensures you know that the document was not cut or truncated due to a file or disk program error:

PATE <eol>

All Publications with Claims have a Claims-Section-Begin Code that implicitly ends the Description section, or ends the Bibliographic section if no Description section exists as is the case with WO Amended Claims files (Kind A4 documents).

Claims Text sections usually begin with code **CLMS**, except for EP documents where the language of the Claims is indicated by one of the three codes **CLMS**, **CLMD** or **CLMF**. EP Applications will have one of these three codes, while EP Granted Patents are unique in that they have Claims in all three languages in the same document with **CLMS** (for English), **CLMD** (for Deutsche or German) and **CLMF** (for French). The Begin Claims codes also have Optional Data which is the Inventor or Applicant *Claims Statement* and is usually a short statement like “**I claim:**” or “**We claim:**” or “**What is claimed is:**”, for example:

CLMS I Claim:

CNTX 1. A system and method to store and process patent related text, including: <eol>
CTXT a series of predefined four character codes, one for each type of data; <eol>
CTXT a single, standardized character encoding scheme shared by all data; and <eol>
CTXT a series of software programs to process 21 different formats converting them into a single, easy to use format allowing both computer programmers and non-computer programmers to make more efficient use of the most popular patent data from around the world. <eol>

CNTX is a numbered claim element in a patent or application, while **CTXT** is and unnumbered Claims Text Element and used for claims where the application author has divided a claim into two or more sections:

CNTX 4. The spot welding method of a high-strength steel plate according to claim 3, comprising the step of, between the step of releasing pressurizing and the step of holding the thin steel plates at the heat treatment temperature after welding, cooling the thin steel plate thereby to reduce a surface temperature of a portion which the welding electrode of the thin steel plate having contacted to a temperature CT (°C) fulfilling a formula (10) below, <eol>
CTXT $CT \leq 150$ (10). <eol>

PBDT is the publication date of the document (all date elements are formatted as YYYYMMDD):

PBDT 20130107 <eol>

PATE is the END PATENT Code and all MAPS files end with this code. This ensures you know that the document was not cut off or truncated due to a corrupted file or disk problem:

PATE <eol>

The Four Logical Sections of MAPS Documents

MAPS files are divided into four Main Logical Sections. The four main sections each contain multiple MAPS data lines; most of which are data elements (lines with a MAPS code and data). Some MAPS codes are grouped together for a specific purpose such as *document identification* or *filing information*. Some MAPS codes have data that may contain more than one data element; such as a document number and a date or a date and a status code (more on this later).

The Bibliographic data section of a MAPS file is **loosely organized** around the order of the elements described in **WIPO Standard ST.9** which is the the “**INID**” (eye-nid) numbering system discussed previously on Page 13. The INID numbers are shown in the outline below in [Square Brackets] after the section names.

Outline Showing MAPS File Organization

The main logical sections and subsections (shown indented) are as follows:

1. Bibliographic Data

- Publication Information (about this document) **[10]**
 - Publication Number [11]
 - Date Published [14]
 - Kind Code of this Publication [13]
 - Language used in Text portions (Description, Claims) [26]
- Correction Information **[15]**
- Application Information (original filing information) **[20]**
 - Application Number (File Number) [21]
 - Date Filed with the Authority [22]
 - Language use with Filing [25]
 - DOADB Family-ID
- Disclaimers, Modifications to Grant
- Priority Claims **[30]**
 - Date, Country and Application Number [31-33]
- Public Availability Dates **[40]**
 - Date of Published Application (first public availability) [41]
 - Deferred Search Report Publication Date [88]
- Technical Group **[50]**
 - International Class [51]
 - International Class [52]
 - Title [54]
 - Patent and Non Patent References [56]
 - Field of Search [58]

- **Related Applications [60]**
 - Various Related Applications [61-64]
 - Previous Publication [65]
 - Reissue Information
- **Concerned/Involved Parties [70]**
 - Applicant [71]
 - Inventor [72]
 - Assignee/Grantee [73]
 - Agent/Legal Rep [74]
- **Data Relating to International Conventions (other than Paris Convention) [80]**
 - Designated States [81]
 - Regional States [84]
 - Begin National Phase Date per PCT Art. 23 or 40 [85]
 - PCT Application Information (in non-WO pubs) [86]
 - PCT Publication Information (pubnum and pubdate) [87]
- **Abstract Text [57]** (intentionally stored last in Bib section – out of numeric order)

2. Description

- Government Interest Statement (US and US sourced PCT)
- Parent Case Text
- Background
 - Field of technology
 - Prior Art
- Summary
- Brief Description of Drawings or Figures
- Details of the Specification

3. Claims and/or Amended Claims

- Optional Statement (I claim... We claim...)
- Numbered Claims
 - Numbered claims are often divided into sectional elements
- Optional Text Statement in reference to Amended Claims (PCT only)
- PCT Kind A4 Pubs are Amended Claims and usually contain the Original Claims in a **CLMS** section followed by the Amended Claims in a CLMA section, both with a set of numbered claims.

4. Search Report Data and Text (when available – currently OCR sourced)

Notes:

- Not all document types have all four main sections
- Not all subsections are included in all documents.
- Purpose of this outline is to show the approximate order of MAPS data.
- Abstract text in MAPS files is the last data item in the Bib section (and not in INID order).

MAPS File Contents by Kind Code

The MAPS file type defines its contents and can be identified by the Kind Code on the Publication Number. The most common type of MAPS files and which of the four sections they contain are listed in the next table. **Numbered Columns are:** 1) Bibliography, 2) Description, 3) Claims, 4) Amended Claims, 5) Search Report.

| Kind | Type | 1 | 2 | 3 | 4 | 5 | Notes: |
|---------|----------------|---|---|---|---|---|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| A | Application | √ | | | | | JP Kind Code 'A' is Unexamined Applications (Abstract Only) |
| A1, A2 | Application | √ | √ | √ | √ | √ | Amended Claims optionally appear with the original claims in WOs. Recent EPs & WOs may have Search Report Text. |
| A3 | Search Report | √ | | | | √ | Only recent EPs have Search Report text. |
| A8 | Bib Correction | √ | | | | | Application Bib Correction |
| A9 | Corrected Copy | √ | √ | √ | √ | √ | Some WO A9s may be corrected Search Reports which will only have Bib and possibly the Search Report data. Correction Code Indicates the contents, see Appendix C for PCT (WO) Correction Codes. |
| A4 | Amended Claims | √ | | | √ | | Pre-2009 WO Amended used Kind Code B1 [see Note 1] |
| B1, B2, | Granted Patent | √ | √ | √ | | | US Grants have an Abstract Text paragraph, EP Grants do not. |
| B3 | Granted Patent | √ | √ | √ | | | US Grants have an Abstract Text paragraph, EP Grants do not. |
| B8 | Bib Correction | √ | | | | | Granted Patent Bib Correction |
| B9 | Corrected Copy | √ | √ | √ | | | Granted Patent Corrected Copy (entire document) |

Note 1: These were generated from DOCDB data since the back-file supplied by WIPO did not contain these documents, and since DOCDB is the source, they also do not contain the Amended Claims text.

Data Types in Bibliographic Fields

MAPS data following each code (beginning in the 5th column) can be a single item, or multiple items depending on which MAPS code it is. Some of the individual elements are described in the following paragraphs.

Dates

Complete dates are always converted to 8 digits (YYYYMMDD) and if it is the only element on the line it is stored directly after the MAPS code and a single space. Partial dates in text form (e.g., "Mar. 1978") found in some fields are usually not converted and left as found. This type of date is often found in patent references.

Country Codes

Along with the rest of the IP industry, we follow the WIPO standard for 2 letter country codes (**ST.3**). We must because most of our source data does, but we still find data where a country name we entered and not the more acceptable two letter code. We convert them where the field is defined as ST.3 compatible. In the EP and WO patents where a field is comprised of one or more country codes (e.g., Designated Contracting States), they can be separated by a space (a space character) or by commas if they contain additional elements with the country code. They will usually be on a single line after the MAPS code that defines them, but they can also be listed in multiple repeating MAPS fields (lines), and this is usually determined by how the data was delivered. We do try to consolidate data to make it easier and faster to parse.

Languages

The Languages specified in the bibliography sections are always **two lower case letters** designated in ISO Publication 639. We have included a subset of ISO 639 in **Appendix B** titled **Language and Country Codes**.

Publication Numbers

Each document published by each Authority is assigned a Publication Number. IP Data Corporation uses the accepted, defacto standard which combines the **Country Code** from WIPO Standard ST.3 that identifies the source patent Authority, the **Sequence Number**, which may also begin with the year published, and the document **Kind Code** to build the complete Publication Number.

Publication Numbers on Corrections Not Unique

The above format is the accepted standard although this does not always allow unique identification of publications in a single collection since correction document are often published with the same number later in the life of an application or granted patent. We add a dash followed by the correction sequence to make them unique when adding them to our database systems. “-2” is the first since the second one is the first to cause a co

US Publication Numbers

The **USPTO (US)** uses several sequences for various kinds of patents, and a year-based sequence that begins at 0000001 (7 digits) following the year for applications. Prior to 2001, all US Kind Codes were a single letter. Utility patents used the letter “A”, and the rest are the same as listed below, minus the digit.

Examples:

| | |
|-----------------|----------------------------------------------|
| USH2532H1 | Statutory Invention Registrations (H prefix) |
| USPP25221P2 | Granted Plant Patents (PP prefix) |
| USRE45324E1 | Reissue Patents (RE prefix) |
| USD720901S1 | Design Patents (D prefix) |
| US9144194B2 | Utility Patents (No letter prefix) |
| US20150100154A2 | US Utility Applications |
| US20150082499P1 | US Plant Applications |

WO Publication Numbers

WIPO (WO) uses a year-based sequence that begins at 000001 (six digits) each year for the World applications (e.g., WO2015000150A1). The WIPO Publication numbers have evolved over the years. The older formats have caused some confusion over the years as they evolved to their current form, and their history is discussed in **Appendix W** titled **WIPO Publication Number Details**.

EP Publication Numbers

The **EPO (EP)** uses a single sequence for both applications and granted patents and the **Kind Codes** change based on the type of publication for the invention. Leading zeros to create 7 digits after the EP are required (e.g., EP0000025A1).

Publication Numbers in EP References

The EPO has a standard format used to list a reference to a publication number in EP documents (mostly patents). It is also found in a few WO and US documents, but not many. We have converted most entries to the standard publication number that can be entered in most search systems around the world. We only do this when we are sure our conversion is absolutely correct. The list below contains examples of most of the different document number we see in the listings, and we also show the converted publication number.

| | | | |
|-------------------|---|---------------------------|---------------------------------------|
| EPO Format | → | Publication Number | |
| CY KC 1 234 567 | | CY1234567KC | (CY is Country Code, KC is Kind Code) |

| | |
|-----------------|------------|
| AU-A- 1 830 195 | AU1830195A |
| EP-A- 0 367 503 | EP0367503A |
| FR-A- 2 658 364 | FR2658364A |
| GB-A- 1 419 750 | GB1419750A |
| US-A- 3 317 539 | US3317539A |

Many of the JP Applications are converted to the new "Renumbered" format using a letter for the Emperor Year (S or H). We do not convert kind codes other than "A" (Applications) since we only have JP Application Abstracts in our inventory, and we can't be sure we are 100% correct on other Kind Codes.

| | | |
|------------------|--------------|-----------------------------------------------------|
| JP-A- 7 250 688 | JPH07250688A | (Added Emperor Year Letter H) |
| JP-A- 50 044 649 | JPS5044649A | (Added Emperor Year Letter S, dropped leading zero) |

| | | |
|-------------------|--------------|----------------------|
| DE-A- 2 102 222 | DE2102222A | |
| DE-C- 456 185 | DE456185C | |
| DE-A1- 2 647 425 | DE2647425A1 | |
| DE-C1- 3 601 220 | DE3601220C1 | |
| DE-U1- 9 211 006 | DE9211006U1 | |
| US-B1- 6 396 864 | US6396864B1 | |
| DE-A1- 36 005 976 | DE36005976A1 | |
| DE-U1- 29 613 910 | DE29613910U1 | |
| SU-A1- 690 543 | SU690543A1 | |
| EP- - 0 527 505 | EP0527505 | EP with NO Kind code |

| | | |
|------------------|---------------|-------------------------------------------|
| WO-A2-91/04547 | WO1991004547A | |
| WO-A-94/10249 | WO1994010249A | |
| WO-A-03/027313 | WO2003027313A | |
| WO-A-20/04037467 | WO2004037467A | |
| WO-A-20/05102565 | WO2005102565A | |
| WO-A-02//01804 | WO2002001804A | Two slant bars are found in a few entries |

Problem children (exceptions)

The next DE would **not** be converted since the information we have does not show this format:

DE-A1-1102004 038 77 ????????????

US Applications require 11 digits to be valid in US inventories, but DOCDB cuts out the '0' after the year. We can correct these since we know where the error was introduced:

US-A1- 2003 183 833 US20030183833A1

JP Publication Numbers

The **JPO (JP)** uses several sequences for the unexamined applications (the English Abstract Collection that we supply). They currently use a four digit year, plus a six digit sequence that begins at 000001 each year, plus they append a single letter Kind Code ("A" for unexamined Applications) which gives us: JpyyyynnnnnnA for all new applications abstracts.

Prior to 2000, the JP applications also used a year-based sequence, but the year was the count of years for the current Japanese Emperor's reign. The "Showa" era (**S**) for Emperor Hirohito ran from 25 December 1926 to 7 January 1989. The Heisei era (**H**) for Emperor Akihito began 8 January 1989 and is still in effect, but the **H** years ended January 1st 2000 when they changed to the current format with the four

digit Gregorian calendar (described above).

Our JP inventory begins October 1976 and run to the currently published week in 2015 (which is 90 days delayed for human translations). The EPO coordinated a major renumbering of all Japanese documents which did the following:

- Added H or S to indicated the Emperor Year in the publications number (not always obvious),
- Forced all Emperor Years to Two Digits,
- Removed all leading Zeros from the Sequence (remaining digits after the 2 digit year).

For our inventory, 1976 Applications begin with JPS51 and run to 1989 (JPS64) which transitions to JPH01, then runs to JPH11 in 1999.

If you see the a Number 1-10034 on a JP patent application image file, the text file will now be numbered JPH0110034. Examples of the new format:

| | |
|---------------|-------------------------------------------------|
| JP2000000125A | New, current style |
| JP2013025997A | New, current style |
| JPH03253A | 3 rd H year - sequence 253 Kind A |
| JPH1011987A | 10 th H year - sequence 11987 Kind A |
| JPS521045A | 52 nd S year - sequence 1045 Kind A |
| JPS622A | 62 nd S year - sequence 2 Kind A |

Conversion of Japanese Emperor Years to Gregorian Years

| | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|
| S51 | S52 | S53 | S54 | S55 | S56 | S57 | S58 | S59 | S60 |
| 1976 | 1977 | 1978 | 1979 | 1980 | 1981 | 1982 | 1983 | 1984 | 1985 |
| | | | | | | | | | |
| S61 | S62 | S63 | S64 | H01 | H02 | H03 | H04 | H05 | H06 |
| 1986 | 1987 | 1988 | 1989 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 |
| | | | | | | | | | |
| H07 | H08 | H09 | H10 | H11 | | | | | |
| 1995 | 1996 | 1997 | 1998 | 1999 | | | | | |

There are JP numbers where the Dash (-) or a space does not exist in the old format, and without the publication date it is impossible to convert it to the new format without research. For example, JP6011025A could be JPS6011025A, or it could be JPH0611025A since leading zeros were not always added when used as a reference in a document. **Keep this in mind when you run across JP number in the old format.**

Defining Multiple Items on One Logical Line

There are two ways to define multiple items of the same MAPS type to be listed in a MAPS file:

- 1) Enter the MAPS code as many times as needed to hold all of the data items, one per line, or
- 2) Separate the items on one line using a space, commas or semicolons (;) between each element.

For larger elements like 50 character Class entries, we use No 1., one per MAPS Code (i.e., one per Line).

If the elements are small, such as a list of two letter Country Codes, these are almost always placed on one line that begins with the MAPS code and separate them with a space. An example of this is the DSTS in all EP data, and DSTS and DSTR in older PCT data, for example:

DSTS AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LU MC NL PL PT RO SE SI

MAPS Codes with Variations

Some MAPS Code may have minor variations to the basic type which is indicated by adding one or two words followed by a colon as the very first elements after the MAPS code. They are same basic data type as the original code, it may simply apply to another region or have a special meaning. All variations are listed with examples in both the MAPS Codes Table and the Quick Reference List of MAPS Codes.

An example of a variation is the Designated Contracting States (**DSTS**) under the EPC called Extended Contracting States. Here is an example of the **DSTS Extended states**:

DSTS Extended states: AL HR LT LV MK YU

Another example are the four Regions listed in the Designated Regional States MAPS code:

DSTR European (EP): AL patent, AT patent, BE patent, ...

DSTR OAPI (OA): BF patent utility-model, BJ patent utility-model, ...

DSTR ARIPO (AP): BW patent utility-model, GH patent utility-model, ...

DSTR Eurasian (EA): AM patent, AZ patent, ...

MAPS Codes with Repeating Sets of Data

In cases where we have many small elements with two or three parts, we can place them all on one MAPS data line and separate the sets with commas. When this is done, the different items will be separated with spaces, and multi-word elements will use a dash to connect the words, which eases parsing. This makes the pattern:

element1 element2 element3, element1 element2 element3, ...

An example of this is the Designated States in **WO** documents that list the type of protection for each country. For example:

DSTR OAPI (OA): BF patent utility-model, BJ patent utility-model, CF patent utility-model, ...

Description Text Section

The Description or Descriptive Text section immediately follows the Bibliography. It begins with the MAPS Code, **DESC**, which contains no data. Like all MAPS codes, it begins in the very first column on the line. The next line is usually an **HDR1** with a standard Section Header describing the contents of the first section, like this:

DESC

HDR1 FIELD OF TECHNOLOGY

DTXT [0001] The present invention relates to an improvement of a buckle ...

HDR1 BACKGROUND ART

DTXT [0002] There is a buckle including a pair of side frame portions, ...

The MAPS code immediately following the **DESC** does not have to be the **HDR1** and in many OCR sourced documents, it is a **DTXT** (a text paragraph).

Section Label Headings Logically Divide Description Text

As shown in the MAPS file sample above, and in the outline further above it, the Description section is usually divided into several logical sections identified by **Text Headings** that describe each section. Unfortunately, the sections and their names are not formally standardized. There are recommended sections in the applications for the different Authorities.

There are now six levels of headings or section headers in the MAPS specification (**HDR1** through **HDR6**). At some point in the future, MAPS Code **HDR1** will be reserved for the major subsection dividers in the descriptive text that are common to most patents and applications from all Authorities. Examples of these include:

Government Interest (US and Some PCTs)
Cross Reference To Related Applications,
Background,
Summary,
Brief Description of the Drawings, and
Detailed Description.

Current data, including the Version 7 Rebuild/Re-Release of all MAPS data (being generated now) will use **HDR1** as it has been used in the past so that very few changes, if any, to parsers of the V6.3f data will be needed.

Since a portion of the data we have is OCR sourced, and section labels are not as standardized as we would like, heuristic algorithms in our software are used to help identify the major sections in many of the OCR sourced documents. In some of the OCR sourced PCT documents where no section headings exist, our editors may add up to 5 standard headings that include:

FIELD OF TECHNOLOGY
BACKGROUND
SUMMARY
BRIEF DESCRIPTION OF THE DRAWINGS, and
DETAILED DESCRIPTION

This is only done if they are absolutely sure they can identify the sections correctly or when XML

Processing Instructions provided by some Authorities positively identify the sections. We have added a comment in the OCR documents where this may happen. **If you intend to index the text for searching, this provides identifiers for the various sections if you want them in different indexes.**

Tables in the Description Section

Tables in MAPS 7.0a files are all plain (UTF-8) text lines built from Green-book and several version of the XML table structures supplied by the USPTO, EPO and WIPO. Our table generation process calculates spacing with a fixed width font, and display in an HTML system requires that you insert a <pre> command before the table, remove the TABL characters as you send each line of data out, then end the table with a </pre> close tag. Parsing each of the words or symbols in MAPS tables is as easy as a text paragraph.

The following two tables were copied directly out of Patent **US7337559B2** published 20080304. Most of the text was removed from the first DTXT paragraph above the first table since it's content was not relevant to the discussion on MAPS tables, but we wanted to show the type of data directly above the table for context.

Patent authorities supply internal numbers for tables in most versions of their bulk data. We usually add this number on the first table line since applicants don't always supply table numbers, as is the case for the first of these two example tables. If the author incorrectly numbered them (or forgot one) the internal number will not match the applicant numbers in later tables, which is what happened on these two tables and also why we chose these tables as our sample table data.

DTXT Data was collected on... Trial information is provided below:

```
TABL      TABLE-US-00001
TABL
TABL      Subject      Trial      Speed
TABL                               (m/s)      Shoe
TABL
TABL      1            1            4.47      Regular
TABL      1            2            4.47      Energy Return
TABL      1            3            4.47      Light Energy Return
TABL      2            1            3.84      Regular
TABL      2            2            4.47      Regular
TABL      2            3            3.84      Light Energy Return
TABL      2            4            4.47      Light Energy Return
TABL
```

DTXT The temporal measure of the running stride were determined to be as follows:

```
TABL      TABLE-US-00002
TABL
TABL
TABL      TABLE 1
TABL
TABL      Temporal Stride Measurements
TABL      Speed      Trial      Stance      Swing      Stride
TABL Subject      (m/s)      Number      Time(s)      Time(s)      Rate(s)
TABL
TABL      1            4.47      1            0.207      0.420      0.627
TABL      1            4.47      2            0.207      0.426      0.633
TABL      1            4.47      3            0.207      0.413      0.620
TABL      2            3.84      1            0.217      0.450      0.667
TABL      2            4.47      2            0.206      0.440      0.647
TABL      2            3.84      3            0.206      0.440      0.647
TABL      2            4.47      4            0.203      0.437      0.640
```

Claims Text Section

The Claims Text section will begin with one of the **Claims-Section-Begin** MAPS codes. The most common is **CLMS**. EP patents contain a copy of the claims in all three of their supported languages. Each has its own Claims Section Begin MAPS code. They are **CLMS** for English, **CLMD** for German, and **CLMF** for French. **CLMS** is also used in PCTs for all of the supported languages (over 30) where the **PBLN** MAPS code (Publication Language) defines the language in the Description and Claims sections. **CLMS** is also in all US documents since they are all in English. One additional code, **CLMA** is for Amended Claims which are only found in **WO** Application documents.

Claims Statement

The **Claims-Section-Begin** MAPS codes optionally have data which is the Claims Statement. The Claims Statement is usually a short statement like “**I claim:**” or “**We claim:**” or “**What is claimed is:**”. Much longer Claims Statements do exist, some as long as 160 characters.

Below is an example of a typical (but fictitious) Claims Text Section:

CLMS I Claim:

CNTX 1. A system and method to store and process patent related text, including: <eol>

CTXT a series of predefined four character codes, one for each type of data; <eol>

CTXT a single, standardized character encoding scheme shared by all data; and <eol>

CTXT a series of software programs to process 21 different formats converting them into a single, easy to use format allowing both computer programmers and non-computer programmers to make more efficient use of the most popular patent data from around the world. <eol>

Dependent Claims

Numbered Claims that continue the description of a part of the invention that builds on and references a prior numbered claim are called Dependent Claims. They use a very practiced wording or language to build on the previous claim, and these include the following phrases:

CNTX 2. The method of claim 1, wherein ...

CNTX 4. The method according to claim 1 or 2 or 3, wherein ...

CNTX 2. A noise reduction circuit as claimed in Claim 1, characterized ...

The next example is another claim that is dependent on the claim before it (claim 3), and it is in two parts because a formula is referenced which is contained in a separate **CTXT** element:

CNTX 4. The spot welding method of a high-strength steel plate according to claim 3, comprising the step of, between the step of releasing pressurizing and the step of holding the thin steel plates at the heat treatment temperature after welding, cooling the thin steel plate thereby to reduce a surface temperature of a portion which the welding electrode of the thin steel plate having contacted to a temperature CT (°C) fulfilling a formula (10) below, <eol>

CTXT $CT \geq 150$ (10). <eol>

Attorneys easily recognize dependent claims as they read and analyze them. Some systems track claims and their structure, including their status as being an independent or a dependent claim. Since we process

and edit a lot of OCR patent data, we attempt to spend a bit more time on the claims text to be sure the OCR process got all of the punctuation correct and that the words claim, claims or claim(s) and the numbers are correct as shown on the legal copy of the document. This allows those parsing data looking for dependent claims to more easily identify them. Some data is shipped with the references in dependent claims marked or tagged in some way, but OCR text is never tagged this way (which is why we spend more time editing the claims text from OCR sources).

Each numbered claim is supposed to begin with a capital letter and should end with a period. Periods should not be used anywhere else in the claims except for abbreviations or “real” numbers in numeric data (e.g., a decimal point). What this says is that each claim can be only one sentence. The attorney’s “artistic talent” is sometimes pushed to the limit to follow this rule and sometimes the sentences look pretty bizarre from the standpoint of grammar, but to the trained eye of professional searchers and attorneys, the sentences usually make perfect sense.

If you are interested in understanding claims structure in more detail, there is an excellent article written by Gene Quinn, an attorney and the founder of **IP Watchdog.com**, here:

<http://www.ipwatchdog.com/2014/07/12/understanding-patent-claims/id=50349/>

Mr. Quinn has a clear, concise writing style, he communicates concepts and ideas very effectively, and it is well worth reading if you have a need to understand why claims are written and formatted the way they are.

Tables in Claims Text

Tables are sometimes contained in claims text. Usually to list a well defined set of ingredients, a list of elements and their relationships directly supporting the claim text, and formatted lists in general. The table is usually the last section in a **CTXT** paragraph, and there should be a period in the last line of the table if it is not followed by another **CTXT** paragraph.

Search Report Text Section

The EPO and WIPO are making changes in the weekly text data to include the Search Report text elements. It is currently included in their weekly EP data, and was included for a while in the PCT data as OCR text. When included, it is located after the Claims Section in A1 Kind documents. Sometimes by itself as the only section after the bibliography in Kind A3 documents.

The EPO has also added XML tags to identify many of the search report elements, but WIPO has not yet done this. We hope that they will use the same tags as the EPO, but this is not always the case.

We have added text label prefixes to allow you to parse the text, if desired. Please keep in mind that it is OCR data. Because of the layout of the Search Reports and the fact that it is usually completed in several different fonts, the OCR is less accurate than the OCR on text that uses a single font.

The Search Report Section MAPS Begin Code (**SRPT**) may also have the text “**Begin International Search Report**” after the Code which simply draws the reader’s (or editor’s) attention to the fact that the Claims have ended if you are in a Kind A1 document. It appears like this:

SRPT Begin International Search Report

The current source of Search Report text is OCR for both EP and WO documents. The EPO indicated that

this should change in the not too distant future, and hopefully WIPO is also headed in that direction.

The Search Report data and the element labels we have added are described in detail in **APPENDIX S** titled **Search Report Text Data Format**.

Uniquely Identifying Publications or Inventions

There are very few unique data fields in the patent text. If you are building a patent database, you will want to ensure that you have a truly unique number to identify your documents. You will also want a method to identify all documents as a group that apply to specific application or invention. These two goals are discussed in the next sections.

Using the Publication Number in a DBS

You **will need to combine the Publication Number and the Publication Date to be sure it is unique**. The reason for this is that WIPO and other patent Authorities publish multiple correction documents with the same Publication numbers. In one of our databases we simply append a unique suffix on the publication number such as (1), (2), (3) etc. when we find one with a different date already in the database. This also shows us how many there are when we look at the latest version of any correction. They also sort as the highest in ascending order.

The Application Number (or File Number) in a DBS

The Application Number (also called the file number) is a unique number assigned by the Authority for each patent application they receive. It is not unique on the documents since it is on every document dealing with the submission of that single application (a single invention). The Application Number is the number that ties all of the documents together that deal with a single application submission. We have seen this referred to as a *Patent Family*. However, there are different definitions for a "Patent Family" and one of the the most popular method of defining a Family are the patents and applications that share a common Priority Application Claim. As always, your patent attorney is your best source for a complete description of the various legal relationships between applications, patents and other legal documents involved in the ever-changing process of patent prosecution.

Variable Length Text Fields

Most sections in a MAPS file are text fields and may contain from one to any number of words where a word is from one (1) to any length of characters long. Note that some indexing systems will only index up to a maximum number of characters (e.g., 48, 65 or 127) of any one word's length. Make sure you know the word length indexing limitations of your system.

There is no set maximum number of words in the line or in a paragraph using any of the MAPS Text codes (ATXT, DTX, CTXT, STXT and CNTX). Also, words have no set maximum length of characters as some data items may be present that are not words.

Many common section titles also have a MAPS code assigned, and the MAPS code will be saved in the MAPS file with no text following it. When the text is built for display from the MAPS file, the MAPS code is usually replaced by the text it represents. This is done only with very common section titles but it saves hundreds of megabytes of storage with over 5 million US text documents, alone (and it also speeds up certain aspects of processing MAPS files).

Primary Languages

Languages in US Documents

All US patent documents are in English, and MAPS **AFLN** (Application Filing Language) and **PBLN** (Publication Language) will always contain “en” (English).

Languages in JP Application Abstracts

The JP Abstracts are translated from the original Japanese applications and the only major text field included is the Abstract translated into English. The **PBLN** field may contain “ja” which indicates the language of the Description and Claims in the original document that are not included in these documents, and NOT the language of the Abstract paragraphs.

Languages in EP Documents

With EP documents, the Title, Abstract, Claims and Description can be in one of three official languages: **English, French or German**. We have two additional MAPS codes for each section to indicate the French or German text. **This is required since all three languages may exist in a single EP document**. This is because EP Patents always have the Claims text in all three languages. The following table shows the sections and the codes used to indicate the language for each section:

| Section | English | French | German |
|---------------------|----------------|---------------|---------------|
| Title: | TITL | TITF | TITD |
| Abstract: | ABST | ABSF | ABSD |
| Claims: | CLMS | CLMF | CLMD |
| Description: | DESC | DESF | DESD |

EP Applications will have the Abstract, Description and Claims in the same language listed in the MAPS **PBLN** field. The Title will be in all three languages, English, German and French. We attempt to acquire English Abstracts for non-English EP applications from several sources, but we do not get all of them, and we do not get them in time to add them to the initial publication. They are added at a alter date and made available later as updated collections.

EP Patents do not have an Abstract paragraph, but they do have the Title in all three languages like the applications. The Description is in the same language that was in the application (as listed in **PBLN**), and all of the claims are contained in all three official languages. If a Search Report section is included, it should be in English except where names or some single elements referenced are in the published language which may be French or German.

Languages in WO Documents

WO documents are delivered in up to 30 different languages and most always include an English and French Title and Abstract along with the Title and Abstract in the primary language of the application listed in MAPS **PBLN**. The French Title and Abstract use same codes used in the EP documents, with two

additional codes for the Title and Abstract in the Primary language shown below:

| Section ↓ | English, | French, | Other (language in PBLN) |
|------------------|----------|---------|--------------------------|
| Title: | TITL, | TITF, | TITO |
| Abstract: | ABST, | ABSF, | ABSO |

The Title and Abstract codes are followed by text in the language indicated by the codes listed in the above table. With PCTs (WO applications) the MAPS fields **TITO** and **ABSO** will contain the title and abstract in the language specified in the MAPS **PBLN** field (Publication Language). The **CLMO** and **DESO** fields are **no longer used since** this specification as of Version 6.3 guarantees that the **DESC** (Description) and **CLMS** (Claims) sections will contain the language listed in the **PBLN** Publication Language field. When the publication language is something other than English, for example, Japanese (**ja**), Korean (**ko**), a Chinese dialect (**zh**) or Russian (**ru**), the Bibliographic data elements will still be in English and also usually supplied in French, and the **TITO** and **ABSO** fields will continue to be used if these data are available (in some instances they are not).

The current ST.9 indicates the language in the field is a two letter language from **International Standard ISO 639:1988**. ISO indicates that the NEW standard is **ISO 639-1:2002**. We checked and none of the abbreviations for languages in the publications we handle have changed so our list in Appendix B of this specification are all accurate. You will find the correct language in the **PBLN** for the publication (not the **AFLN** field). Also, don't look at the country of origin that is part of the document number or from an application number and assume, for example, that finding JP means the text is in Japanese (**ja**). Quite often it isn't. We are now receiving **Japanese PCTs in English** that are NOT OCR sourced.

All MAPS files in all languages are supplied encoded in UTF-8 binary character codes, and the Full Text Sections (Description and Claims) are in the language specified in the MAPS code **PBLN** (e.g., **PBLN de** for German). The paragraph and table indicators used inside of the text sections remain the same (ATXT, DTXT, CTXT, CNTX, TABL, etc.) no matter what language is indicated for the section.

All current WIPO Standards may be found here: http://www.wipo.int/standards/en/part_03_standards.html

DOCUMENT TYPES SUPPORTED IN MAPS VERSION 7

The following table lists abbreviations for **document types** supported by MAPS. They are used in **Column 2 of the MAPS CODE TABLE** (below) to indicate which document types use the described MAPS codes. The abbreviation “Bib.” is for Bibliographic Data often called “Front Page Data” on a patent or application.

| | |
|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ALL | All types use the code where ALL is listed (USA, USB, EPA, EPB, WO, JPA). |
| USA | US Patent Applications (first published in March 2001) |
| USB | US Granted Patents |
| EPA | European Patent Applications |
| EPB | European Granted Patents |
| WO | WIPO or PCT (Applications filed under the Patent Cooperation Treaty) |
| JPA | Japanese Patent Application (Bib Data w/Abstract text paragraph) |
| Notes: | <p>1) These are not MAPS codes. They are abbreviated document types to indicate which document types use the MAPS code described in each table entry.</p> <p>2) Document type abbreviations for all <i>Special Order</i> Abstract format other than the JPA have been removed for clarity. Any of the full-text publication collections may be converted to “Abstract” versions which include the Front Page “Bib” elements and the Abstract Text Paragraph.</p> |

MAPS CODE Table

Examples of **MAPS** and **MAPS-XML** document sections from one or more current documents are usually included after each section of the table. Even though the MAPS and MAPS-XML documents contain the same data and share the MAPS codes, you will note minor differences where a MAPS code is used as Group tag in the XML instead of directly containing data, for obvious reasons. It makes the XML a bit “cleaner” and easier to parse since XML is handled as a serial stream of characters (or bytes) instead of in a line-oriented fashion, which is how MAPS files are designed to be parsed.

Columns In the MAPS Code Table

The MAPS Code table rows are divided into Groups of MAPS codes by purpose that follow the INID Table Groups. The MAPS Code Table is further divided into the **six columns** that include:

1. **MAPS Code**
2. **Document Types** that use this MAPS Code (EPA, EPB, USA, USB, WO and JPA)
3. **Element Type** – One or more letters, each indicating a characteristic of the data defined and follows the MAPS code on the same line after a single space character (described below in detail).
4. **INID/ST.32 code** – See **Table of INID Codes on page 13**. May also contain an ST.32 data tag to help further define the data for those familiar with these codes.
5. **Example(s)** of the code with data, if it fits in the call width, otherwise examples will follow at the end of the Section or Group in table.
6. **Comments and/or Details** about the code in addition to information found elsewhere.

Element Types

Element Types are letters that describe an elements characteristics that may help database designers define fields or storage definitions for the data they will be loading. These letters are defined next and are contained in **Column 3** in the table with the heading **Type**. One or more of the letters will be used for each MAPS code defined.

- **D – Data Element(s)** – One or more elements exists with a predefined format and order as part of the data. This may be an application or document number, date, country, status or a repeating group of two or more single data elements, as defined for that particular MAPS code. Unformatted text may follow defined formatted data elements with a separator, if specified.
- **U – Unique** – This field alone or when combined with a second field (as described) may be used as a Unique element in a database for a collection, or all collections, if so described (**U** implies **S**).
- **R – Repeating** – More than one line with this MAPS code containing an element or a set of elements (if defined) may appear in a MAPS file, or, multiple data elements or sets (if defined) may appear on one MAPS line with this code. This does not mean there *will* be more than one, only that there *may* be more than one.
- **M – Mandatory** – At least one of this MAPS code with any defined accompanying data will exist in the type of MAPS file(s) listed in column 2. Any exception will be described in column 5 or 6.
- **S – Single entry** – indicates that if this MAPS code exists, it will not repeat in the MAPS file (or tag Group in MAPS-XML) and only a single copy of this code and the defined data will exist in any one document (or tag Group for MAPS-XML). If it *must* exist, **M** will be added, and if it is Unique, **U** will be added.
- **O – Optional Elements** – One or more elements in the data is optional and may not be included with the data, as specified in the description.

- **T – Text** – All or a portion of the data following this maps code on the line should be treated as text information, and may consist of, from one character up to a very large paragraph (e.g., 50 KB). When Data and Text are combined (**DT**), The text data always follows the defined formatted last and follows the defined data element(s), on the line.
- **L – Multiple versions exist for different languages** – Indicates that another MAPS code exists with the same definition as this code, but it is for a different human language.

Document Identification Group [10]

Known as the **Document Reference Section** in DOCDB and other document data schema, this section uniquely identifies the document and some of its characteristics such as its name, kind, publication date and the primary language of the text sections of the document's contents.

The MAPS code for the Publication Number is always the first code in a MAPS file and identifies the document type and Authority. The publishing Authority is further identified by the two letter country code that always begins the publication number, and the specific document purpose is indicated by the Kind Code that ends the Publication number. Along with this code is the Issue or Publication Date, Kind Code, Document Name Label, Publication Language and MAPS Version (new for V7.0a). Plus, if the document is a "Further Publication" such as a correction or delayed Search Report, you may also find a code with the Publication Date of the original application (**APPD**), as described below.

| MAPS CODE | Docs used in | Type | INID ST.32 | Example(s) of MAPS code and Data (if any) | Description of MAPS Code and Data Included with it (if any) |
|----------------------------|--------------|----------------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Publication Numbers | | | | | |
| PATA | USA | U ⁽¹⁾ D M | [11] | PATA US20150167112A2 PATA US20110030112P1 A2 -Utility, P1 -Plant Application CAUTION: DOCDB Only uses 6 of the 7 digit sequence for US Apps. | US Application Publication Number Country, 4 digit Year, 7 Digit Sequence, 2 char Kind Code (1 letter, 1 number). See Appendix X for Kind Code Descriptions for all Patent Authorities in our collections |
| PATB | USB | U ⁽¹⁾ D M | [11] | PATB US8137811B1 (Utility) PATB USD690123S1 (Design) PATB USRE49812E1 (Reissue) PATB USPP20115P2 (Plant) PATB USH1276H1 (a SIR pub) | US Granted Patent Publication Number Though not a patent, Statutory Invention Registrations (SIRs) also use PATB since they publish with the Grants. |
| USAA | USAA | U ⁽¹⁾ D M | [11] | USAA US20130010891A1 US Abstract Collections are generated on demand (Special Order). | US Application Abstract. A US Application minus the Description and Claims. MAPS PATE will immediately follow ABST . |
| USPA | USBA | U ⁽¹⁾ D M | [11] | USPA US8001254B2 U.S. Abstract Collections are generated on demand (Special Order). | US Granted Patent Abstract. A US Patent minus the Description and Claims text sections. MAPS PATE follows ABST or ATXT . |
| PATW | WO | U ⁽¹⁾ D M | [11] | PATW W02015001847A1 This format is used on all PATW WO publications in our collection, but some references to WO pubs may still use older formats. * | WO Publication Number. PATW is used on all WO Applications and Further Publications under the Patent Cooperation Treaty (PCT). * See Appendix WO for an in depth discussion on how PCT Publication numbers have changed over time. |

| | | | | | |
|-------------|--------------------|----------------------------|------|-------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| PATU | EPA EPB | U ⁽¹⁾ D M | [11] | PATU EP1770823A1 PATU EP2331013B1 PATU EP1770823A1 | EP Publication Number. PATU is used on all EP documents. Kind Code (PKCD) further determines type. EP Pubnums are always 7 digits with a Kind 2 Character Kind Code. |
| PAJA | JPA | U ⁽¹⁾ D M | [11] | PAJA JP09H1109A PAJA JP2007011013A | Publication Number for JP Abstract. Kind Code (PKCD) further determines type. |
| PNAL | ALL | U ⁽¹⁾ D | [11] | PATW W01983004463A1 PNAL W08304463A1 (old vers.) | Publication Number Alias. May be added where original Publication Number format has changed. |

⁽¹⁾ The **Publication Number** combined with the **Issue Date (ISSD)** or **Publication Date (PBDT)** in the same document (always on the very next line or in the same Group in MAPS-XML), forms a **Unique Key** to identify the document within the single collection, and from all collections from that Authority, and even from all collections from all Authorities worldwide as long as those authorities follow the same format for their Publication Numbers (**Format:** Country, optional-Year, Sequence Number, Kind Code).

Publication Dates

| | | | | | |
|-------------|-------------------------------------|-------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ISSD | USB EPB* | S D M | [45] B450 | ISSD 201301014 | Issue Date for a US Patent. Serves the same purpose as PBDT, except it also indicates date the patent was Granted. Format: YYYYMMDD (* EPBs should now uses PBDT since further publications will not have same meaning, but ISSD is still valid in First Published Grant) |
| PBDT | EPA PCT* USA USB | S D M | [14] B140 | PBDT 20140107 * Note: ALL PCT MAPS Publications now used PBDT correctly. APPD is added in Further Publications to indicate publication date of Original Application. | Date Published - Date that the document which contains this code was published. Format: YYYYMMDD |
| APPD | USA EPA WO | S D M | [43] B430 | Example from Delayed ISR: PATW W02013101350A3 PBDT 20150108 APPD 20130704 ... May also appear later in Bib in some pubs (same data, same meaning) | Date of Publication of First Application (A1 or A2 Doc Kinds). Used in Further Publications or Corrections to identify the Initial/First Public access as a printed application. Do not confuse this w/ Date Filed. |

Publication Kind Codes

| | | | | | |
|-------------|------------|-------------|------|---------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| PKCD | ALL | S D M | [13] | PKCD A1 PKCD B8 See Appendix K titled Kind Codes, All Authorities for details | Kind Code - One or two characters. Usually 1 letter, or a letter plus 1 digit. Indicates document “ kind ” (purpose) assigned by the authority. Examples: A, A1, A2, A3, A4, A8, A9, B1, B2, B3, B8 B9, S, S1, E, E1, P, P1, P2, P3, P4, H, H1, etc. |
|-------------|------------|-------------|------|---------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Publication Language

| | | | | | |
|--------------------------------------------|-------------------------------------------|----------------------|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| PBLN | USA USB EPA EPB WO | S D M | [26] | PBLN en PBLN ko PBLN de PBLN zh NOTE: On JP English Abstracts, this indicates the language in the Description and Claims of the complete published Application. | Publishing Language - This is two lower case letters from ISO-639 and is the language of the Description, Claims and Search Report Text sections. EP A and B publications have special language requirements for content. See MAPS codes: TITL, TITD and TITF for Titles, ABST, ABSD and ABSF for Abstracts, DESC, DESD and DESF for Description, and CLMS, CLMD and CLMF for Claims. |
| Document Label (Name) | | | | | |
| LABL | ALL | R* M | | LABL United States Patent | Document Name Label – The first line is the primary document name. * One or two optional lines may be added that modifies or enhances the name and description in some way. |
| MAPS Version and File Creation Date | | | | | |
| MVER | ALL | S D M | | MVER V7.0a-20150923 | MAPS version and File Creation Date (yyyymmdd). If two copies exist of a publication, always use the file with the latest MVER date. |

Document Identification Section MAPS Examples

PATB US8621662B2
 ISSD 20140107
 PKCD B2
 PBLN en
 LABL United States Patent
 MVER V7.0a-20150923
 ...

PATW W02013062354A9
 PBDT 20150205
 APPD 20130502
 PKCD A9
 PBLN ko
 LABL INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)
 LABL Published: corrected copy with international search report
 MVER V7.0a-20150923
 ...

Document Identification Section MAPS-XML Example

```
<bib>
  <pub-data>
    <pubnum>US8621662B2</pubnum>
```

```

<pbdt><date>20140107</date></pbdt>
<pkcd>B2</pkcd>
<pbIn>en</pbIn>
<labl>United States Patent</labl>
<mver>V7.0a-20151103</mver>
</pub-data>
...
</bib>

```

Correction Information [15]

In EP and PCT documents, a Kind Code of A8 indicates a front page correction to the original A1 or A2 Kind document with the same sequence, and A9 usually indicates a complete corrected copy. However, even though the Kind A9 indicates a complete “corrected copy” WIPO documents are a bit more complicated because the A9 may be a corrected copy of the original application, or a corrected copy of the Search Report. The Correction Code is the key to identifying exactly what the correction contains. PCT corrections usually contain the current correction codes and date as well as all past corrections including Codes and Dates (in order from current down to the oldest). **APPENDIX C** titled **Correction Codes for PCT (WO) Pubs** describes PCT Correction Codes in detail.

| MAPS CODE | Docs used in | Type | INID ST.32 | Example(s) of MAPS code with Data (if any) | Description of the MAPS Code and Data |
|-------------|---------------------------|-------------|------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CORR | EPA EPB WO | R D T | [15] B152EP B665WO | See Examples after the table., Information about this corrected copy. For B8 (Bib correction) affected INID field will be listed. For text correction (kind B9) affected page range may be listed. | Correction Information EPs indicate the date, correction sequence (1, 2, 3, etc), the name of the section corrected, (Bibliography, Claims, etc.) or the INID number for Front Page corrections in Kind A8 or B8, type of correction and may also contain the Bulletin number where 1st published PCTs will contain Publishing Sequence (3, 4, etc), Correction Type (R4, R5, R6, R7, etc.), the document number and any text included with the correction information. |
| CODT | EPA EPB | R D | [48] | CODT 20150205 PCTs may also have Bulletin No.: CODT 19980123 Bulletin: 199804 Only found on corrected copies. | Date of Correction - This is the date of a correction document (the one in which it is listed, or a past correction). Format: YYYYMMDD Bulletin: YYYYNN |
| DSRP | WO EPA EPB | S D | [88] | DSRP 20150205 Typical DSRP entry in EP Kind B (Grant) pub: DSRP 20111005, Bulletin: 201140 | Date of Search Report – The date the Search Report was Published. This may appear in EP-B documents with ONLY the date and a Bulletin Number. as well as in the A3 (Delayed) Search Reports, themselves. |

Correction Information Examples

from MAPS **WO2013062354A9**:

CORR 4, R6, Bulletin: 05/2015, Change to the Translation of the international search report due to an error during technical preparations for international publication.

CODT 20150205

CORR 3, R4, Bulletin: 28/2013

DSRP 20130711

from MAPS-XML **WO2013062354A9:**

```
<corr>
  <text>4, R6, Bulletin: 05/2015, Change to the Translation of the international
    search report due to an error during technical preparations for
    international publication.</text>
  <date>20150205</date>
</corr>
```

from MAPS **EP2770399A9:**

CORR W1, Description

CODT 20151021 Bulletin: 201543

```
<corr>
  <text>W1, Bulletin: 201543, Description</text>
  <date>20151021</date>
</corr>
```

- Notes:**
- 1)** W1 – The digit (after W) is the sequence number of this correction for this publication. W2 would indicate that this is the second correction made to this publication.
 - 2)** **CORR** and **CODT** are combined to make the **<corr>** group in MAPS-XML.
 - 3)** Last item on the line (or in the **<text>** element) is which section is being corrected (one or more)

Application Information Group [20]

This contains information on the **Original Application** (*the actual Filing*) and contains the File Number, also called the Application Number, assigned by the Authority the date the application was filed, the Language it was filed in, and the DOCDB Family ID.

| MAPS CODE | Docs used in | Type | INID ST.32 | Example(s) of MAPS code and Data (if any) | Description of MAPS Code and Data Included with it (if any) |
|--------------------------------------------------|--------------------------------------|--------------|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Application (File) Number | | | | | |
| APNM | USA USB | S D M | [21] | APNM 223015 APNM 000020 This used to be the Serial Number defined in US Applications as SNUM . US Apps now use APNM . | Application Number (US, 6 digit) - A complete US Application number or File Number uses the 2 digit SCOD + 6 digit APNM providing an 8 Digit File Number (uniquely defining an application for an invention). Use APNO for NON-US pubs. Future US MAPS Pubs may use APNO if the USPTO abandons the SCOD . (We keep it because they still do). |
| APNO | WO EPA EPB | S D M | [21] | EP: APNO 02078334.6 WO: APNO PCT/ZA2011/000052 | Original Application Number - Assigned by the authority (EP, WO or JP). US may use this MAPS code in the future for complete 8 digit File number. |
| SCOD | USA USB | S D* M | [21] | SCOD 09 * - The combined SCOD and APNM fields are unique in US Publications. All 8 digits are required to access the PAIR system as the USPTO. | Series Code – A 2 digit number, but may also appear as letter “D” on some Design patents. Design SCOD is now 29 . SCOD+APNM = complete US file number often with slant bar after SCOD: NN/123456. |
| Application Date | | | | | |
| APDT | ALL | S D M | [22] B220 | APDT 20150106 | DATE FILED - Application filing date. Format YYYYMMDD |
| Application Filing Language | | | | | |
| AFLN | USA USB EPA EPB WO JP | S M | [25] | AFLN zh NOTE: On JP English Abstracts , this indicates the language in in the Description and Claims of the complete Application. | Filing Language - This is 2 (lower case) letters from ISO-639 which represents the language the applications used in the original filing of the application. |
| Application Processing Language (EP only) | | | | | |

| | | | | | |
|-------------------------------|--------------------|---|--------|----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------|
| PRLN | EPA EPB | S | [25] | PRLN de | Processing or Procedure Language - Appears in many EP documents and is the language used to process the application. |
| Family ID (from DOCDB) | | | | | |
| FMID | ALL | S | [30] * | FMID 48168096 * Not part of INID 30, but created from INID 30 priority data. | Family ID - A 4 digit integer defined by DOCDB assigned to applications and patents sharing a common Priority Claim. |

Application Information Group Examples:

From MAPS **US8621662B2**:

SCOD 13
APNM 175987
FMID 45999631
APDT 20110705
AFLN en

From MAPS-XML **US8621662B2**:

Note: The **SCOD** and **APNM** from MAPS have been combined to form a “unique” US Application Number as **<apno>** which is the format required to locate the file prosecution history in the USPTO PAIR system.

```
<appl-data>
  <apno>13/175987</apno>
  <apdt><date>20110705</date></apdt>
  <afln>en</afln>
  <fmid>45999631</fmid>
</appl-data>
```

from MAPS **EP1145820B1**:

APNO 01400849.4
APDT 20010403
AFLN fr
PRLN fr
FMID 8848796

NOTES: Many of the MAPS Codes in **EP MAPS documents** have been moved into the MAPS “preferred order” that is loosely based the INID numbering. The remaining tags that are still out of order together instead of in the Publication and Application Information Sections. We have added and corrected codes to meet the V7.0 spec, but we have not yet re-ordered the output. Newer MAPS file generation software will take care of this issue. This should not affect parsing the data. **MAPS-XML files** are already grouped and reordered correctly.

Priority Application Group [30]

Priority applications are entries that list a legally related application (the file number), country where filed, and the file date (*not in that order – the reverse, in fact*). The listed application affects certain legal dates of the application in the current document, and may affect the length of time this application will be afforded

protection, if granted. Exactly how and why, and for how long is where patent attorneys earn their pay with hard work (*Get a good one. There are lots of good ones out there!* Just another way to say **it is beyond the scope of this specification**). Meanwhile, what do the letters **FADL** have to do with the word Priority? In the APS system, FADL was the **Foreign Application Data Listing**, and we kept it. The three elements and how they are listed in the MAPS data are described below.

| Priority Application Data Entry | | | | | |
|----------------------------------------|------------|----|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| FADL | ALL | RD | [27] [30] | CMNT Date Ctry Appl. Number FADL 20110711 JP 2011-153169 FADL Entry form an older Grant: FADL 20010420 [JP] 2001-122178 Some entries may have [brackets] around the Country. These are going away, but some may creep back in if we have to rebuild files for some reason before we finish some program updates. | Priority Application Listing yyyyymmdd, [2 letter country code], then Application Number from that country. An older data source had text dates and the brackets, but we think all of them have been converted – Please let us know. Example: FADL Apr 20, 2001 [JP] 2001-122178 |
| CMNT | ALL | RT | N/A | The comment line above the FADL entry is simply to identify the items to the Human readers – a convenience line. | In older data this may have been FAPD which stood for Foreign Application Priority Data. |

Priority Application Group Examples

From MAPS File **WO1990004492A3** Published **19900503**

```

CMNT Date      Ctry      Appl. Number
FADL 19881026    GB      8825063
FADL 19890125    GB      8901586
    
```

From MAPS File **WO2012049285A1** Published **20120419**

```

CMNT Date      Ctry      Appnum
FADL 20101014    US    61/344,806
FADL 20101014    EP    10187595.3
    
```

From MAPS File **EP1451952B1** Published **20120419** (w/ MAPS-XML below it)

```

CMNT Date      Ctry      Appnum
FADL 20011109    [US]    337472 P
    
```

```

<fad1>
  <date>20011109</date>
  <ctry>US</ctry>
  <apno>337472 P</apno>      (the EPO adds the P for U.S. Provisional Applications which is SCOD 60 or 61)
</fad1>
    
```

NOTES: The Application Numbers (File Numbers) are not standardized in any way and are kept as the national applicants entered them since requests for subject matter or related applications from each nation's archive usually requires the file number in the original format.

Public Availability Dates Group [40]

EP documents include additional information on various legal dates indicating public availability during the various stages of prosecution from the first public availability of the application through issuing the patent, and some events after the grant on to expiration. Not all dates are included in all EP patents, but far more than most other Authorities. Your patent attorney can assist with the importance of each during database development.

| | | | | | |
|------|------------------|----|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| AMDT | EPB | S | [45] B477 | AMDT 20140115 Bulletin: 201403 | Date of grant and mention of the opposition decision Date followed by a Gazette/Bulletin number |
| APPD | EPA EPB WO | O | [43] B430 | APPD 20001220 Bulletin: 200051 | Date of 1st Published Application May also be in Publication Info Group in Further Publications. 1 st Public Availability |
| INTG | EPB | SD | B451EP | INTG 20011212 | Date of Announcement of Intention to Grant Without Preceding Examination Report YYYYMMDD (Date only) |
| NOIG | EPB | SD | B452EP | NOIG 20041124 | Notice of Intention to Grant Date YYYYMMDD (Date only) |
| B241 | EPB WO | SD | | B241 20000814 | Examination Request Date Not published on face of EP or PCT |
| B242 | EPB WO | SD | | B242 20031031 | Dispatch Date of 1st Exam. Report Not published on face of EP or PCT |
| B243 | EPB | SD | | B243 20140115 | Date of Patent Maintained as Amended Not published on face of the EP |
| B244 | EPA WO | SD | | B244 20060513 GB | Date of Request for Conversion to National Application Contains Date and Country |
| B245 | EPA EPB | SD | | B245 20100226 Susp-indicator: 5 New data will use: YYYYMMDD Susp-indicator: 5 Data before V7 may have: YYYYMMDD Susp. Indicator: 5 or may contain only suspension indicator without the date (an error, now corrected). Not published on face of EP or PCT | Date of suspension (Rule 13) / interruption (Rule 90) of proceedings Suspension Indicator according to EPC rules: 1 Suspension (Rule 13) 2 Applicant died (Rule 90(1)(a)) 3 Legal incapacity of applicant (Rule 90(1)(a)) 4 Legal prevention (Rule 90(1)(b)) 5 Representative died (Rule 90(1)(c)) 6 Legal incapacity of representative (Rule 90(1)(c)) |
| B246 | EPA EPB | SD | B246 | B246 20090216 | Date of Resumption of Proceedings Not published on face of EP or PCT |
| B248 | EPA EPB | SD | B248 | B248 20130815 | Date of Notification Rights after Appeal Not published on face of EP or PCT |
| B405 | EPA EPB WO | SD | B405 | B405 20140115 Bulletin: 201403 | Date and Bulletin Number announcing first availability of this publication. Date <i>should</i> match PBDT. Format: YYYYMMDD Bulletin: YYYYNN |

| | | | | | |
|-------------|--------------------|----|----------------|------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| B450 | EPB | SD | [45] B450 | B450 20020717 Bulletin: 200229 | Date of publication and mention of the grant of patent (EP B1) or Mention of the grant of patent (EP B2) Contains Date and bulletin No. This is date the B1 published and was mentioned in the Bulletin (Found on EP B1/B2 Pubs). |
| B453 | EPB | SD | [45] B453EP | B453 1) B3, 20140101 Bulletin: 201401 | Date of publication and mention of the limitation decision On Kind EP B3 Pubs only. Contains sequence (1, 2, 3, etc.) of published decision and Kind Code (B3) followed by date and bulletin number. Format: 1) B3, YYYYMMDD Bulletin: YYYYNN |
| B473 | USB EPB | SD | | B473 20030117 | Disclaimer Date Standard Date (YYYYMMDD) plus Bulletin No. in EPB. |
| B47D | USB | ST | B473US | B47D This patent is subject to a terminal disclaimer. | This patent is subject to a terminal disclaimer The above text is always included as data. |
| B474 | USB EPB | SD | B474 | B474 14 | Term of Grant Rarely appears in primary EPB documents. Most often used in US Design Patents for the 14 year term. Always whole years in US patents. |
| B47T | USB | SD | B474US | B47T 1628 B47T 5 Years | Patent Term Extension Will always be: “5 years” or will be an integer 1 to 4 digits long which is the number of days of extension. The word “Days” will not appear in the field. |
| B475 | EPB | RD | B475 | B475 20110502 FR; 20100825 NL B475 replaces TRMG used in older MAPS documents. | Lapse of Patent Contains Date and Country. May have multiple entries per MAPS code, or multiple MAPS codes with one date and country. |

Technical Information Group [50]

Title, classifications, cited patents and non-patent literature, and technical field(s) searched are all included in the Technical Information.

Title Section

Titles in US documents are in English, only. But EP documents can be in one (the submitted language), two or all three of the EPO approved languages. WO documents always have the Titles in English, French, and optionally in a third language, which is one of up to 30 different languages. The languages of the title match the codes as described below, with code TITO (Title, Other) indicating it is in the language found in **PBLN**, which is the same language as the Description and Claims.

| MAPS CODE | Docs used in | Type | INID ST.32 | Example(s) of MAPS code and Data (if any) | Description of MAPS Code and Data Included with it (if any) |
|-----------|------------------|------|--------------|----------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| TITL | ALL | STL | [54] B541 | TITL CCD Image Array | Title in <i>English</i> |
| TITF | EPA EPB WO | STL | [54] B542 | TITF CCD matrice d'image | Title in <i>French</i> . EP Applications will be one of 3 languages (en,1 de, fr). WO apps usually have en and fr titles and abstracts. |
| TITD | EPA EPB | STL | [54] B542 | TITD CCD-Bild-Array | Title in <i>German</i> . EP Grants can contain all 3 EPC approved languages (en, de, fr) |
| TITO | WO | STL | [54] | TITO 一种抽屉式蒸笼 A CMNT line may indicate the language but is not required since PBLN specifies the language. | Title using language specified in PBLN . This can be one of over 30 languages supported by WIPO under the PCT. |

Examples of the MAPS Titles are in the descriptions, above.

Classification Section

Inventions are usually classified under two systems; International Patent Classifications (IPC) which is a system maintained by WIPO, as well as “**National Class**” entries from the classification system(s) used by the national publishing authority. The US has now moved to the Cooperative Patent Class (CPC) system which is also the primary system for the EPO (*i.e.*, all EP documents). JP Abstracts in English and PCT (WO) documents are primarily classified under the IPC, but the EPO is adding CPC classifications on both collections, as time permits (and, we applaud the effort and we also send you a BIG, HEARTY **THANKS!!**).

IMPORTANT CONCEPT: Each time IP Data Corporation generates new collections for a new or updated MAPS specification, we completely replace the older IPC entries (V2 through V7) with IPCR (Reclass V8 Data) and National Class entries from the latest Master Classification Files (MCF) or DOCDB, if it contains newer data. **However, users should be aware that classifications may change as corrections are made, or as reclass orders are issued by the authorities for a variety of reasons** (why this happens is beyond the scope of this document). We place updated US, EP, WO and JP classification data in our CSV format on our FTP site for our subscribers usually on the first week of each month. See the “**ReadMe**” text files on the FTP site for current information.

| MAPS CODE | Docs used in | Type | INID ST.32 | Example(s) of MAPS code and Data (if any – and space permitting) | Description of MAPS Code and Data Included with it (if any) |
|----------------------------|--------------|------|------------|------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------|
| CPC Classifications | | | | | |
| CPCV | ALL | SD | [52] | CPCV 20130101 | CPC Version - Optional entry if CPC edition date is not listed as part of the entry. CPC version is a date: yyyyymmdd |
| CPCP | ALL | SD | [52] | CPCP B60L 11/02 20130101CLI20150106BH | CPC Primary Class - Follows ST.8 Standard, but not exactly (a slight change was made late 2013) |

| MAPS CODE | Docs used in | Type | INID ST.32 | Example(s) of MAPS code and Data (if any – and space permitting) | Description of MAPS Code and Data Included with it (if any) |
|-----------|--------------|------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| CPCI | ALL | RD | [52] | CPCI B60L 11/03 20130101CLI20150106BH | Class, Invention related - second or subsequent classification assigned. |
| CPCN | ALL | RD | [52] | CPCN B60L2200/36 20130101CLA20150106BH | Class, Non-Invention related (or in addition to the invention) Zero or more per pub See Discussion later in this spec. |
| CPCC | ALL | RD | [52] | CPCC B29C 66/1122 20130101CLI20150106BH ,1,1 CPCC B29C 65/00 20130101CLI20150106BH ,1,2 CPCC B29C 66/43 20130101CLI20150106BH ,2,1 CPCC B29C 65/00 20130101CLI20150106BH ,2,2 Combo entries are 50 chars like other CPC entries, plus Group and Rank numbers | CPC Combination Classes The entry has the ST.8 formatted class followed by Group and Rank Number: CPCC ST.8-Class, Group #, Rank # |

US Classifications

| | | | | | |
|------|------------|----|------|---------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CLSU | USA USB | SD | [52] | CLSU 53/282 | US Class/Subclass, Primary entry IMPORTANT: The USPTO is no longer classifying Utility patents in the US Classification system. CPC is it! |
| CRCU | USA USB | RD | [52] | CRCU 53/281 CRCU 53/283 CRCU 53/504 CRCU 53/55 | US Class/Subclass, Cross-Reference One CRCU entry for each cross reference classifications assigned (e.g., 1 per line). 0 or more entries. |

IPC Classifications

| | | | | | |
|------|-----|----|--------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CICE | ALL | SD | B516 | CICE 7 CICE 20060101 Date versions may also be included in parenthesis after CLSI class entries. | International Class Edition A single digit, usually 2 through 7, is the IPC Edition for IPC entries in the document. Editions past 7 are an 8 digit date. 20060101 is current edition. |
| CLSI | ALL | SD | [51] B511 | CLSI B60L 9/00 20060101AFI20150106BH Same entry as above using an abbreviated format (when all data is not provided): CLSI B60L 9/00 (20060101) Some spaces have been removed in 1 st example above. All newer IPC format entries include all ST.8 50 characters w/ spaces as placeholders. | International Class entry (Primary) One line per IPC Entry. Abbreviated CPC entries may have the Edition date in parens after the shorter 16 character CPC class entry (example): B41J 2/175 (YYYYMMDD) |
| CCRI | ALL | RD | [51] B512 B513 B514 B515 | CCRI B60L 11/02 20060101ALI20150106BH Same entry as above using an abbreviated format (when all data is not provided): CLSI B60L 11/02 (20060101) | Additional IPC Classifications Abbreviated CPC entries may have the Edition date in parens after the shorter 16 character CPC class entry. For example: B41J 2/175 (YYYYMMDD) May also have leading zeros to fill 4 spaces in Main Group, ONLY: B41J0002/175 |

Field of Classification Search

| MAPS CODE | Docs used in | Type | INID ST.32 | Example(s) of MAPS code and Data (if any – and space permitting) | Description of MAPS Code and Data Included with it (if any) |
|-----------|--------------|------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| FOSS | USB EPB * | RD | [58] | <p>FOSS CPC: B60K 6/48, B60K 6/00, B60L 11/14, B60W 10/026, B60W 10/08, B60W 20/00, B60W 10/02, B60W 10/06</p> <p>FOSS USC: 701/22, 192/30R, 192/31, 192/54.1</p> <p>*Only in EPB when search report data is included (or found in A3 document).</p> | <p>Field of Search – Technology fields identified by class for initial national or international search.</p> <p>Class system is listed prior to first FOSS entry (as UPC, CPC and IPC):</p> <p>FOSS USC: 2/65, 2/67 FOSS CPC: A47K 13/165 FOSS IPC: A47K 13/165</p> |

Classification Section Examples:

From **USD720885S1** US Design Patent:

CPCP A45D 8/02 20130101CFI20150106BH
 CLSI 2803 (Locarno Class Ed. 10)
 CLSU D28/39

Note: Most US Design patents do not have CPC classifications which is why the USPTO continues to maintain that portion of the US Class system.

From **US8978174B2** US Utility Patent:

CLSI A61H 33/04 20060101AFI20150317BH
 CCRI B05B 7/04 20060101ALI20150317BH
 CCRI A61H 33/02 20060101ALI20150317BH
 CPCP B05B 7/0425 20130101CFI20150317BH
 CPCI A61H 33/6052 20130101CLI20150317BH
 CPCN A61H2033/021 20130101CLA20150317BH
 CPCN A61H2033/022 20130101CLA20150317BH
 CLSU 4/541.6

Detailed information on the ST.8 Classification format is contain in **Appendix T** titled **ST.8 Standard Classifications**, and you can always go to the ST.8 standard for the official information but it does NOT mention the minor changes for the CPC. Our discussion is aimed at the basic 50 character format of the ST.8 standard.

Patent and Non-Patent Cited References Section

Cited References are included in documents published **after an official search has been conducted** by a search authority. Some publishing Authorities include who cited the reference such as **cited by examiner**, or **cited by other** often found in US patents.

IMPORTANT NOTES:

1) MAPS codes for section begin/heading text in the Cited References sections are being phased out over 2016. Methods and designs for patent text display systems are so diverse, there is no point

telling users what is on the face of one Authority's printed patent. MAPS codes **REFS**, **URFS**, **FRFS** and **ORFS** that are still left in some MAPS files are being removed from new data files generated over 2016.

2) MAPS codes **UREF** (US References) and **FREF** (Non-US References) both in US Patents will be replaced with **REFP** (Patent Reference) mid to late calendar year 2016. Older Green Book did not have the US country code or Kind codes listed with the US references, which is why they were needed.

3) The “cited by other” and “cited by examiner” will be upgraded to add hyphens in between the words (“cited-by-others”, etc.) for standardization of parsing over the year 2016. Look for it both way until year's end.

| MAPS CODE | Docs used in | Type | INID ST.32 | Example(s) of MAPS code and Data (if any) | Description of MAPS Code and Data Included with it (if any) |
|-----------|--------------|------|------------|-------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| REFS | USB EPB | S | [56] | Heading with no data See Important Note 1 , above | “ References Cited ” heading text. Heading text for obvious sections are being phased out over 2016. |
| URFS | USB | S | [56] | Heading with no data See Important Note 1 , above | “ U.S. References ” (in US Pubs) |
| FRFS | USB | S | [56] | Heading with no data See Important Note 1 , above | “ Foreign References ” (in US Pubs) |
| ORFS | USB EPB | S | [56] | Heading with no data See Important Note 1 , above | “ Other References ” |
| | | | | | |
| REFP | ALL | RD | [56] | REFP US5551212A | Patent Reference - a patent reference from any authority. Always begins with Publication Number; Optionally has Pub Year and Month, Inventor, and Category (who cited it) |
| UREF | USB | RD | [56] | UREF US5551212A See Important Note 1 , above | US Patent Ref – lists publication number, and optionally: inventor’s last name, Full or partial date (Text: “Apr 2009” or Numeric 20090400), Cited By: Examiner or Other |
| FREF | USB | RD | [56] | FREF GB1133461A See Important Note 1 , above | Non-US (foreign) document number references in US docs |
| OREF | USB EPB | RT | [56] | OREF | Non Patent Literature (NPL) Refs include books, magazine articles, conference papers, Internet URLs, etc. |

US Patent and Non-Patent Cited Reference Examples

From US Patent **US8112831B2**

UREF US6155015A Dec. 2000 Kirby cited by other
 UREF US6175971B1 Jan. 2001 O'Neil cited by other

Related Documents Group [60]

This Group contains entries that list applications (by **Application Number** *also known as File Number*) that are “**related to**” the subject document (the one they are found in), as well as other MAPS codes that contain document numbers that are related in some way to the subject document. The greatest number of entries are Related Applications that are mostly Parent applications that are in some way the source of the application for the subject document and provide continuity leading to the sources of the applications. Code documenting Reissue patent (REIS and some RELA codes) and codes for prior published versions (RELP) of the subject document are also included. The following **RELA** relationships are possible, but not all of them exist in the data: An **additional** application, a **divisional** application, **continuations**, **reissues**, **divisional-reissues**, **reexamination**, **merged-reissue-reexamination**, **substitute**, and **provisional** applications.

Related Applications Section

| RELA | USA USB | R D T | | See examples below table. Possible Relationships | Related Application Mandatory Elements: Relation AppNum AppDate [optional items] Optional Items: Parent-status: Abandoned/Granted/Pending Granted-as: US Patent Number (digits only) Parent-PCT: PCT PubNum Child: AppNum US Relationship Abbreviations: CON – Continuation or Continuing CIP – Continuation In Part DIV – Division of ADD – Addition to SUB – Substitution for 371 – a-371-of-international |
|------|------------|-------------|------|--------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| | | | [61] | ADD-Parent | |
| | | | [62] | DIV-Parent | |
| | | | [63] | CON-Parent CON-REISSUE-Parent CIP-Parent 371-Parent (2001-2004, Note 1) | |
| | | | [64] | DIV-REISSUE-Parent REISSUE-Parent REEXAM-Parent REEXAM-3rdParty-Parent (Note 2) Reexam-Reissue-Merger-Parent | |
| | | | [66] | SUB-Parent (Substitution-for) | |
| | | | [68] | Provisional | |

Related Application Examples

RELA CON-Parent 10075921 20071026 Granted-as: 7113245 Child: 09031249

```
<rela>
  <relation>CON-Parent</relation>
  <apno>10075921</apno>
  <apdt><date>20071026</date></apdt>
  <text>Granted-as: 7113245 Child: 09031249</text>
</rela>
```

RELA DIV-Parent 10252728 20020923 Granted-as: 6766565 Child: 10840718

```
<rela>
  <relation>DIV-Parent</relation>
  <apno>10252728</apno>
  <apdt><date>20020923</date></apdt>
  <text>Granted-as: 6766565 Child: 10840718</text>
```

</rela>

RELA CON-Parent PCT/CH02/00663 20021204 Parent-Status: PENDING Child: 10860328

RELA CON-Parent PCT/EP2004/000786 20040129 Parent-Status: PENDING Child: 11196080

RELA DIV-Parent 09936640 20020409 Parent-Status: PENDING Parent-PCT: PCT/US00/40039
20000316 Child: 10896602

<rela>

<relation>DIV-Parent</relation>

<apno>09936640</apno>

<apdt><date>20020409</date></apdt>

<text>Parent-Status: PENDING Parent-PCT: PCT/US00/40039 20000316 Child: 10896602</text>

</rela>

RELA REISSUE-Parent 08476198 19950607 Granted-as: 05635235 Child: 09/325122

Notes:

1. In older US application data (2001 to 2004) a related application relationship of <a-371-of-international> that we list as **371-Parent**, was removed in the data and the PCT document number is now listed with the parallel US application as the <parent-pct-document> as part of the <parent> group. So when a PCT was the source of the US application, the additional field will appear in the <text> area.
2. REEXAM-Parent is a Reexamination requested by the Applicant, Assignee or the Commissioner, while the 2nd option (REEXAM-3rdParty-Parent) is someone other than the Applicant, Assignee or the Commissioner.
3. Most Application Numbers have the Series Code and Serial Number which provides 8 digits, but some of the older RELA entries did not include the SCOD (Series Code) and only have the 6 digits Serial Number in the entry. All of them have the File Date which provides the necessary data to lookup the SCOD, if needed.

| RELA | EPA EPB JPA | R D T | | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|-------------|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| | | | B610 [61] | ADD-Parent – earlier applications to which this App. is added , |
| | | | B620 [62] | DIV-Parent – earlier applications from which this App. was divided , |
| | | | B620EP [62] | DIV-Parent-EP – Child listed was divided from the application for this document (this is a special EP relationship not fully defined in the DTD/XLS specs) |
| | | | B621EP [62] | B621EP-Application -special relation to EP Appl. (the source of this App.) |
| | | | B622EP | B622EP-Publication -a special relation to a Published EP document (also the source) [B622EP] |
| | | | B630 [63] | CON-Parent -an application from which this application a continuation [B630] |
| | | | B631 [63] | CON-Parent -an application from which this application a continuation [B631] |
| | | | B632 [63] | CIP-Parent -an Appl. from which this Appl. is a continuation-in-part [B632] |
| | | | B633 | CON-Parent-Reissue -a reissued application from which this application is a continuation [B633] |
| | | | [64] B640 | Reissue-Parent -an Appl. reissued from which this Appl. is the source [B640] |
| | | | [64] B645 | Reexam-Parent -an Appl. reexamined from which this application is the source [B645] |
| | | | [66] B660 | SUB-Parent -an application from which this application is derived and is a substitute [B660] |
| | | | [67] B670 | Utility-Basis-Parent - Number and filing date of a patent application, or number of a granted patent, on which the present utility model application or registration is based - [B670] |
| <p>NOTES on EP and JPA RELA entries:</p> <p>1) The order of EP RELA elements is the same as the US (INID for INID)</p> <p>2) A status for the parent is rarely supplied in the EPs.</p> <p>3) The “Child Application Number” is rarely supplied in EPs, but when it is, the first RELA entry is always (or should be) the application number of the publication in which it is listed, (i.e., indicating “this is the child of the listed parent” - which begins the continuity of the parent relationship).</p> <p>4) JP RELA entries are always DIV-Parent relationships. Older SGML sources did not contain the relationship and the entries only have the application number. Newer versions will have the DIV-Parent relationship plus the application number. Dates are not supplied by the JPO for the Related Applications.</p> | | | | |

From MAPS file **EP2119355B1**:

RELA DIV-Parent 08160661.8 20080717 Granted-as: EP2025228

From MAPS-XML file **EP2119355B1**:

```
<rela>
  <relation>DIV-Parent</relation>
  <apno>08160661.8</apno>
  <apdt><date>20080717</date></apdt>
  <text>Granted-as: EP2025228</text>
</rela>
```

From MAPS file **JP2015157814A**:

RELA DIV-Parent 2011527902

From MAPS-XML **JP2015157814A**:

```
<rela>
  <relation>DIV-Parent</relation>
  <apno>2011527902</apno>
</rela>
```

| | | | | | |
|-------------|------------|---|------|-------------------------------------------------------------------------------------------|---------------------------------------------------------------------|
| RELP | ALL | R | [65] | RELP US20120278971A1 20121108 Usually the Published Application for a US Patent | Prior Publication Number Earlier version of this document |
|-------------|------------|---|------|-------------------------------------------------------------------------------------------|---------------------------------------------------------------------|

The **REIS** code contains information about the Reissue patents sequence (1 to n) provided in an older format of the US Data known as Green Book data. With the exception of the Sequence number, the same data is now provided in the **REISSUE-Parent** relationship of the **RELA** MAPS code [INID 64] or ST.32 B640 series data.

| | | | | | |
|-------------|------------|---|------|---------------------------------------|-----------------------------------------------------------------------------------------------------|
| REIS | USB | R | [64] | REIS examples are listed below | Reissue Information - MAPS Reissue patents prior to 2001 will contain REIS data elements. |
|-------------|------------|---|------|---------------------------------------|-----------------------------------------------------------------------------------------------------|

REIS Reissue (1st) of Appnum: 001353 Filed: 19900820 Patnum: 4210046 Issued: 19910828
REIS Reissue (2nd) of Appnum: 234567 Filed: 19950730 Patnum: 5551212 Issued: 19960828
REIS Reissue (3rd) of Appnum: 700121 Filed: 19950730 Patnum: 5551212 Issued: 19960828

If Sequence is not listed in the source data, the following format will be found in the REIS data:
REIS Reissue of Appnum: 123456 Filed: 19950730 Patnum: 5551212 Issued: 19960828

Notes on data item in REIS line:

- Appnum:** the application number of the patent being reissued
- Filed:** the date the Appnum was filed in the format **YYYYMMDD**
- Patnum:** The U.S. patent number being reissued
- Issued:** The date **Patnum:** was issued (the Publication date)

Concerned Parties Group [70]

Concerned parties include the **applicant**, **inventors**, the **agent** (*legal representative*) and **assignee** (*grantee*) for all publication Kinds, plus we have **opponents** for EP patents. The US is currently the only Authority to provide the Patent Examiner and sometimes the Assistant Examiner, and we keep them with this group, as well. We always define the Extended Fields (with X in the MAPS code) if at all possible.
Appendix A, titled Name and Address for Concerned Parties

Opponents and Our IID Database

Opponents were added to EP data after the EPO Version 1.4 as XML was placed in service (~2011). It was quite a few months before any real Opponent data began to show up, and nearing the end of 2015, we still only see only 20 to 30 files, each week with opponent data.

We initially had problems accurately dividing the data to provide more granularity, and this has not been solved with brute force (heuristic algorithms), however, PLAN B is looking good. We have been building our own EPO "IID" DBS that we will share with our subscribers the first or second quarter of 2016. This DBS allows us to both verify data provided by the EPO, and complete missing pieces of information when the IID is in our database. We are now seeing a 65% hit rate in our tests on the OPON data, so it is approaching the point of being truly useful. When this happens, we will announce two new MAPS Codes and generate the 20 to 30 additional EP documents in parallel to the new 7.0a and ask for user comments. Until this happens the **OPON** field will continue to hold the Opponent and Agent fields exactly as they have been provided for the last 4 years or so. The sections are in the OPON table and plenty of examples are provided.

The EPO provides too few fields and the attorneys don't seem to be satisfied providing a reduced set of data. This is why we can find a very long law firm names in the <snm> (sir name) field, and in the same opponent data line we see an attorney's name (last, first and middle), a PO Box, a street name, **and** a long law firm name, all between a pair of <str> (street) tags. It is difficult to parse accurately. They also include Line Feeds after each section in an attempt to delineate the different items in the over-stuffed field to make it appear like a mailing address (the EPO also does this with Japanese names and addresses). And, it almost works, which is like saying "my parachute **ALMOST OPENED** before I hit the ground!" (we're fairly certain that those words are rarely uttered, after the fact).

See the Notes after the Opponent Table and Examples near the end of the Concerned Parties Group for more information.

Applicant Section

For for details., See **Appendix A** titled **Name and Address for Concerned Parties**

| | | | | | |
|-------------|------------|----|------|-------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| APLX | ALL | R | [71] | APLX Smith; Frederic M. Midland, TX US With extended data, when available. 1 logical/physical line per inventor. | Applicant - 5 sub-fields provide granularity for address, city, country, etc., One inventor per line. See Appendix A for Extended Party Field details. |
| APPL | USA USB | RT | [71] | APPL Smith; R.L. (Conroe, TX, US) *US law prior to AIA required that the Inventory be the applicant. | Applicant – Single text line, loose formatting, used ONLY when Applicant data can't be parsed and categorized. APPL may also have more then one Applicant per line. The Inventor data may appear on one line or multiple MAPS lines. |

Inventor Section

For details., See **Appendix A** titled **Name and Address for Concerned Parties**

| | | | | | |
|-------------|------------|----|----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| INVX | ALL | R | [72] [75] | INVX Smith; Frederic M. Midland, TX US With extended data, when available. 1 logical/physical line per inventor. | Inventor - 5 sub-fields provide granularity for address, city, country, etc., One inventor per line. See Appendix A for details. |
| INVS | USA USB | RT | [72]* [75]* | INVS Smith; R.L. (Conroe, TX, US) *US law prior to AIA required Inventor also be Applicant which is INID [75]. Inventor Only is INID [72]. | Inventor – Single text line, loose formatting, used ONLY when Applicant data can't be parsed and categorized. APPL may also have more then one Applicant per line. The Inventor data may appear on one line or multiple MAPS lines. |

Assignee Section

For details, see **Appendix A** titled **Name and Address for Concerned Parties**

| | | | | | |
|-------------|-----|----|------|-----------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ASGX | ALL | RD | [73] | ASGX Sony Corporation Tokyo JP [03] With extended data, when available. 1 Assignee per line | Assignee - 5 sub-fields provide granularity for address, city, country, etc., One Assignee per line. Role is 2 digits [00] to [09] See Appendix A for Extended Party Field details. |
| ASSS | ALL | RT | [73] | ASSS Sony Corporation (Tokyo,JP) | Applicant – Single text line, loose formatting, used ONLY when Applicant data can't be parsed and categorized. APPL may also have more then one Applicant per line. The Inventor data may appear on one line or multiple MAPS lines. |

Agent Section

For for details., See **Appendix A** titled **Name and Address for Concerned Parties**

| | | | | | |
|-------------|-----|----|------|----------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| AGTX | ALL | RD | [74] | ASGX Kenyon & Kenyon, PLLC US With extended data, when available. 1 Agent per line | Agent - 5 sub-fields provide granularity for address, city, country, etc., One Agent per line. See Appendix A for Extended Party Field details. |
| AAFM | ALL | RT | [74] | AAFM Kenyon & Kenyon PLLC, US Some Trivia AAFM – Agent, Attorney and Firm was the Legal Rep field in the Green Book. | Agent – Single text line, loose formatting, used ONLY when Applicant data can't be parsed and categorized. APPL may also have more then one Applicant per line. The Inventor data may appear on one line or multiple MAPS lines. |

From MAPS Patent **US8990974B2**:

APLX Thomas; Kurt J. ||Indianapolis, IN|US| [--,US] (app-type: applicant, desig: us-only)
APLX Brown; Derek A. ||Lizton, IN|US| [--,US] (app-type: applicant, desig: us-only)
APLX Enlow; Brian A. ||Noblesville, IN|US| [--,US] (app-type: applicant, desig: us-only)
INVX Thomas; Kurt J. ||Indianapolis, IN|US| (designation: us-only)
INVX Brown; Derek A. ||Lizton, IN|US| (designation: us-only)
INVX Enlow; Brian A. ||Noblesville, IN|US| (designation: us-only)
AGTX Faegre Baker Daniels LLP|||--| (attorney)
ASGX Masco Corporation of Indiana||Indianapolis, IN|US| [02]
ASGX Mercury Plastics, Inc.||Middlefield, OH|US| [02]
PEXM Crane; Lauren, Dept. 3751

NOTES: The Added text data after the Party fields shown in the examples is discussed in detail in **Appendix A** titled **Name and Address for Concerned Parties**.

| OPON | EPB | R | [78] | Opponent | Opponent and Agent |
|------|-----|---|------|----------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| | | | | Information on organization/persons with active opposition to grant. | OPON data consists of: Opponent sequence no. 01, 02, etc), Organization Name [IID: 123456789], P.O. Box No. 0000, Street Address, post-code, City, Country, Rep: Firm or attorney name [IID: 123456789], P.O. Box No. 0000, Street Address, post-code City, Country |

From MAPS **EP121278B2**:

OPON 01, 20051201, Akzo Nobel Coatings International N.V. [IID: 100699610],
 Velperweg 76, 6824 BM Arnhem NL, Rep: De Vries, Adriaan Jacobus
 [IID: 100758764] , Akzo Nobel N.V. P.O. Box 9300, Velperweg 76,
 6800 SB Arnhem NL

OPON 02, 20051201, Consortium für Elektrochemische Industrie GmbH [IID: 100700790],
 Zielstattstrasse, 20, D-81379 München DE, Rep: Fritz, Helmut, et al
 [IID: 100034263], Wacker Chemie AG, Zentralbereich Patente, Marken und Lizenzen,
 Hanns-Seidel-Platz 4, 81737 München DE

OPON 01, 20110310, Haldex Brake Products GmbH [IID: 100702192], Mittelgewannweg 27,
 69123 Heidelberg DE, Rep: REHBERG HÜPPE + PARTNER, et al [IID: 100060390],
 Patentanwälte PartG mbB, Robert-Gernhardt-Platz 1, 37073 Göttingen DE

OPON 02, 20110328, Knorr-Bremse, Systeme für Nutzfahrzeuge GmbH [IID: 100702966],
 Moosacher Str. 80, 80809 München DE, Rep: Specht, Peter, et al [IID: 101103861],
 Loesenbeck - Specht - Dantz, Patent und Rechtsanwälte, Am Zwinger 2, 33602 Bielefeld DE

Examiner Section

The US provides Primary Examiner, and sometimes the Assistant Examiner.

Examiner Section

| | | | | | |
|-------------|------------|----|--|-----------------------------------|------------------------------------------------------------------|
| PEXM | USB | ST | | PEXM Polito; Nicholas, Dept. 3673 | Primary Examiner Primary Examiner's Name and Dept. No. |
| AEXM | USB | ST | | AEXM Hare; David R | Assistant Examiner Assistant Examiner's Name |

PCT and National Phase Group [80]

This group covers PCT application information, and more specific information about these applications as they move through the National Phase as countries prosecute the application moving towards granting a patent or some other protection document.

Designation States Section

| | | | | | |
|-------------|--------------|----|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DSTS | WO EP | RD | [81] B810 | Designated States EP and PCT Designated Contracting States (per ST.3) EP lists states only, while PCT may also contain Types of Protection (by document name) for each state. See format discussed after DSTR entry. | Later listings of PCT Designated States usually contain types of protection by document type and the types are listed on a document named: “Types of Protection Available via the PCT in PCT Contracting States” (Search the WIPO web site for this title) |
| DSTR | WO | RD | [84] B840 B845EP | PCT patent regions include: DSTR European (EP): DSTR OAPI (OA): DSTR ARIPO (AP): DSTR Eurasian (EA): The two letters are State Codes in ST.3 assigned to the Regions. See Examples Below | Designated Regional States List of states within defined PCT defined cooperative patenting regions. In early documents, these were often combined in one section with national designated states. Designated Extended States will also be listed here. In later WO pubs, one or more protection types described in the next table row. |

Format of DSTS and DSTR entry with protection type:

DSTS AU patent, AZ patent utility-model, BA consensual-patent patent, BB...

Note that **commas** separate each country while **spaces** separate the country and its protection type(s). Protection types of more than one word are hyphenated to become a single entity for easier parsing of multiple types. The list of Protection Types is updated about once a year by WIPO. The most common Protection Types include:
consensual-patent inventor-certificate patent utility-certificate utility-innovation utility-model

Search the Internet or WIPO for **“Types of Protection Available via the PCT in PCT Contracting States”** to find the current versions. If we find new types from future lists, we hyphenate multiple word names, and remove apostrophes. All other letters are preserved.

From MAPS file **WO2012232656A2**:

DSTR European (EP): AT patent, BE patent, BG patent, CH patent, CY patent, CZ patent, ...
DSTR OAPI (OA): BF patent utility-model, BJ patent utility-model, ...
DSTR ARIPO (AP): BW patent utility-model, GH patent utility-model, ...
DSTR Eurasian (EA): AM patent, AZ patent, BY patent, KG patent, KZ patent, ...
DSTS AE patent utility-model, AG patent, AL patent utility-model, ...

From MAPS-XML file **WO2012232656A2**:

```

<dsts>
  <ctry>AE</ctry><ptype>patent</ptype><ptype>utility-model</ptype>
  <ctry>AG</ctry><ptype>patent</ptype>
  <ctry>AL</ctry><ptype>patent</ptype><ptype>utility-model</ptype>
  <ctry>AM</ctry><ptype>patent</ptype><ptype>provisional-patent</ptype><ptype>utility-model</ptype>
  ...
  <ctry>ZW</ctry><ptype>patent</ptype>
</dsts>
<dstr>
  <region>EP</region>
  <ctry>AT</ctry><ptype>patent</ptype>
  <ctry>BE</ctry><ptype>patent</ptype>
  <ctry>BG</ctry><ptype>patent</ptype>
  <ctry>CH</ctry><ptype>patent</ptype>
  ...
  <ctry>TR</ctry><ptype>patent</ptype>
</dstr>
<dstr>
  <region>OA</region>
  <ctry>BF</ctry><ptype>patent</ptype><ptype>utility-model</ptype>
  <ctry>BJ</ctry><ptype>patent</ptype><ptype>utility-model</ptype>
  <ctry>CF</ctry><ptype>patent</ptype><ptype>utility-model</ptype>
  ...
  <ctry>TG</ctry><ptype>patent</ptype><ptype>utility-model</ptype>
</dstr>
<dstr>
  <region>AP</region>
  <ctry>BW</ctry><ptype>patent</ptype><ptype>utility-model</ptype>
  <ctry>GH</ctry><ptype>patent</ptype><ptype>utility-model</ptype>
  <ctry>GM</ctry><ptype>patent</ptype><ptype>utility-model</ptype>
  ...
  <ctry>ZW</ctry><ptype>patent</ptype><ptype>utility-model</ptype>
</dstr>
<dstr>
  <region>EA</region>
  <ctry>AM</ctry><ptype>patent</ptype>
  <ctry>AZ</ctry><ptype>patent</ptype>
  <ctry>BY</ctry><ptype>patent</ptype>
  <ctry>KG</ctry><ptype>patent</ptype>
  ...
  <ctry>TM</ctry><ptype>patent</ptype>
</dstr>

```

International Filing Information Section

Patent Cooperation Treaty (PCT) Applications

“Native” PCT applications are those filed with Authorities who accept them as their own and don't require their nation's application to be submitted with the PCT application. The EPO accepts about 1000 native PCT applications each week where no EP-A is published, but they do publish the abbreviated bibliographic “stub”. We include them with the collections and the weekly subscription data. The USPTO (U.S. law) requires a US application to be submitted with the PCT application. WIPO maintains a PCT database online, here: <https://patentscope.wipo.int/search/en/search.jsf>

Hague Agreement

The **PCT does not support Design patents** which is where the **Hague Agreement** comes in. Like the PCT, the Hague Agreement provides a standardized application for Designs, and WIPO maintains the data. The data elements are similar, except the Hague Agreement provides for a “**Registry**” and we end up with a single Registration Number instead of an Application and Publication Numbers. The Publication Date we receive is first public availability of the Design application (via WIPO's online database, we assume). The WIPO Design database is located here: <http://www.wipo.int/designdb/hague/en/>

The following four data elements are provided by the USPTO for US Design filings registered under the Hague Agreement, and the first column shows the existing MAPS code used for each element:

| MAPS Code | Hague Data from USPTO Bulk data | Example data |
|-----------|---------------------------------------------|--------------|
| IPNO | International registration number | DM/086482 |
| PPDT | International registration publication date | 20150605 |
| PCFD | International filing date | 20150513 |
| P102 | International registration date | 20150513 |

The PCT data for International Utility-type applications provided in US, EP and JP publications are the **Application Numbers** (also called **File numbers**) and sometimes the Publication number. While the Hague Design system **provides ONLY a Registration number**. Since both the PCT and the Hague are International Filings for Applications, we share the MAPS elements with identical definitions for the two systems to reduce programming requirements. Only the MAPS **P102** is re-purposed since it was not populated since 2003, but still defined, and was a date field used for a similar function (We hope this reduces programming requirements for MAPS users).

International Number Formats

The EPO provides the DOCDB format for **PCT Application Numbers** they provide in bulk patent data. It is the first example listed in their eSpace DBS (e.g., **DK2010000021**). The USPTO prefers WIPO's application number format (**PCT/DK2010/000021** or **PCT/DK98/00006**). We are standardizing PCT application numbers on the WIPO format, and we have tested it in the eSpace system and it is properly recognized (*we test them because we use them all the time for research*). We also standardize the PCT Publication Numbers in the MAPS data (**WO2013000123A1**) since the newer format is valid in all source Authority systems. The kind code is not always included. If you validate PCT Publication Numbers in the IPNO field, you may have to make a minor change in your parser to allow for the added Hague Registration numbers, **Format: DM/ + 6 digits (DM/086482)**. We have seen the format: “DM/086 482” but we remove any spaces, if provided.

PCT Publication and Application Number Section

| | | | | | |
|-------------|--------------------------------------------|----|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| IPNO | EPA EPB USA USB JPA | SD | [87] | IPNO W02003012345 Older W0 formats are discussed in Appendix W. Kind Code is rarely supplied. IPNO DM/086482 Hague Design Registration | International Publication Number - PCT Publ. No. - Found in EP, US and JP pubs, this is the published PCT that parallels the national application or replaces it as the native application in the case of EuroPCTs. Hague Design Reg. NO. – International Reg. Number for US Design applications under the Hague Agreement |
| PPNM | USB | | | Discontinued - New Pubs must use MAPS Code IPNO | PCT PUB. NO. - Use IPNO Still appears in V6.3 and older files. |
| PPDT | US EP | SD | | PPDT 20071213 or PPDT 20071213, Bulletin: 200750 | PCT or Hague Design Publication Date – The date the PCT application is published or first public availability of Design Registration under the Hague Agreement. EPs May also have bulletin Number where announced. |
| PCNO | USA USB EPA EPB | SD | [86] | PCNO PCT/DK2010/000021 Some EPs may contain language: cc after the application number and a comma. | PCT Application Number New V7.0a EPs should use the PCT/ format with the optional language after the appnum. PCNO PCT/Appnum, Language: en |
| PCFD | USA USB | SD | [86] | PCFD 20100219 PCFD 20100219, Lang: ja | PCT or Hague Design File Date – Date the PCT Application or Design Application was filed with the Authority |
| P102 | USB | SD | B864 | P102 20150513 P102 20010104 | US Design Hague Registration Date , or PCT 102(e) Date in pre-2003 US publications where P102 referred to a date in 35 USC Section 102(e) and was discontinued in 2002, now re-purposed for Hague Registration Date in May 2016. |
| P371 | USB | SD | [86] B863 | P371 20080818 | PCT 371 Date - Exact meaning of P371 date has changed to follow changes in US law over last 12 years. Currently, P371 refers to US XML date field <us-371c124-date> which refers to 35 USC section 371(c)(1), (2) and (4) . |
| P372 | USB | SD | [86] B86? | P372 20150611 | PCT 372 Date - New for 2015, P372 refers to US XML date field <us-371c124-date> which refers to 35 USC section 371(c)(1) and (2) . |

Search Report Date Section

| | | | | | |
|-------------|---------------------------|----|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DSRP | EPA EPB WO | RD | [88] | DSRP 20111005, Bulletin: 201140 The EP MAPS-XML <DSRP> may be in the PCT filing section or in the Correction section. Both are in the same scope. | Date of Search Report – EP-B (Grants) usually have the Date of the search report and the Bulletin in which it was announced which is not related to a CORR entry (like used in PCT documents). |
|-------------|---------------------------|----|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Example from MAPS File **US8991055B2**:

PCNO PCT/IB2007/000862
 PCFD 20070322
 P371 20081201
 IPNO W02007141605
 PPDT 20071213

Example from MAPS file **EP2373095B1**:

PCNO PCT/CN2009/071640, Language: zh (see Note 1 below)
 PCFD 20090505
 IPNO W02010072063
 PPDT 20100701
 DSRP 20111005, Bulletin: 201140

Example from MAPS-XML file **EP2373095B1**:

```
<dsrp>
  <date>20111005</date>
  <text>Bulletin: 201140</text>
</dsrp>

<pcno>
  <apno>PCT/CN2009/071640</apno>
  <pcfd><date>20090505</date></pcfd>
  <text>Language: zh</text>
</pcno>
<ipno>
  <pubnum>W02010072063</pubnum>
  <ppdt><date>20100701</date></ppdt>
</ipno>
```

Example from US Design Patent **USD754922S1**:

```
<pcno>
  <apno>N/A-Hague-Agreement</apno>
  <pcfd><date>20150513</date></pcfd>
  <p102><date>20150513</date></p102>
</pcno>
<ipno>
  <pubnum>DM/086482</pubnum>
  <ppdt><date>20150605</date></ppdt>
</ipno>
```

NOTES:

1) In EP documents, the EPO includes the language in which the application was filed. We add it after the File Number along with the label **Language:** The language is two lower case letters from ISO 639, also listed in **Appendix B** titled **Country and Language Codes** in this specification.

2) **Of Interest:** The **Hague Agreement Registration Number** appeared to be synonymous with a “**Publication Number**” when discussed being “published” in their database which is why we store it in **IPNO** and not **PCNO**, **plus**, if it served the same function as a File Number, the Registration Date (**P102**) would ALWAYS match the Date Filed (PCFD), **BUT**, our research shows that approximately 1 out of 1000 do NOT match. Why not, was not obvious from our research. As always, we recommend consulting your attorney.

Abstract Text Group

The Abstract is usually the last element in the bibliographic data before the Description section begins. We informally consider it part of the Bib section, though technically it is not bibliographic data. The Abstract is defined as a single text paragraph that concisely describes the invention. A small table, a formula or an additional paragraph or two will sometimes be included.

| MAPS CODE | Docs used in | Type | INID ST.32 | Example(s) of MAPS code and Data (if any, length permitting) | Description of MAPS Code and Data |
|-----------|------------------|------|------------|-----------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|
| ABST | ALL | ST | [57] | ABST followed by abstract text paragraph in the English language | Abstract text paragraph in English. |
| ABSD | EPA EPB | ST | [57] | ABSD Abstract text paragraph in German. | Abstract text in German (de) |
| ABSF | EPA EPB WO | ST | [57] | ABSF Résumé paragraphe de texte | Abstract in French (fr). WO Apps almost always have titles and abstracts in both English and French |
| ABSO | WO | ST | [57] | ABSO - followed by abstract text paragraph in language specified in MAPS code PBLN . | Abstract text paragraph in WO that is in a language found in PBLN . |
| ATXT | ALL | ST | [57] | ATXT This is an additional abstract text paragraph. A single paragraph is highly recommended. | Follow-on paragraph of text in language of previous Abstract MAPS Code. Must follow ABSx MAPS code or Table Code that follows ABSx. |
| TABL | ALL | RT | [57] | TABL – A line in a table in the Abstract section. | |

Description Group

This is the main text section that is the description of the invention in an application that becomes the Patent Specification, if granted.

95% of all applications are written in a format that contains five basic sections that include:

- Field of technology
- Background
- Summary
- Brief Description of the Drawings (or Figures)
- Detailed Description (of one or more embodiments)

Additionally, some application may also have one or more paragraphs before the Field of Technology text begins that describes related applications. And, one more paragraph may precede the Field of Technology in US applications and patents that describes any interest the government may have in the invention. Because of inconsistencies in the WO OCR text, and in the US application and patent text, we no longer treat the Government Interest and Related Applications paragraphs as separate sections of the documents. They are now treated as another paragraph in the Description section, usually with their own heading that describes the paragraph's purpose.

Description Section and General Text

| | | | | | |
|-------------|--------------------|-----|--|-----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| DESC | ALL | SL | | DESC A label is on a single line by itself | Begin Description in Language Specified in MAPS Code PBLN |
| DESF | EPA EPB | RTL | | DESF In the French Language | Begin Description in French |
| DESD | EPA EPB | RTL | | DESD In the German Language | Begin Description in German |
| DTXT | ALL | RT | | DTXT This is a paragraph. | |
| TABL | ALL | RT | | TABL Spaces are significant (keep them!) | Table Text Line |
| HDR1 | ALL | RT | | HDR1 Heading Text (Level 1) | Heading Text Line entry HDR2 through HDR6 are defined but not in use. They will be added later in 2016. |
| CMNT | ALL | RT | | CMNT Comments are not part of the specification and are usually inserted to provide additional information for the user processing the file. | Comment Line |

Example of a typical Description Section with all lines truncated (...) except for the headings, and a few line in each area removed (... ..)

From MAPS file for **PATB US8621667B2**:

DESC

DTXT The present invention is an item of headgear comprised of a ...

HDR1 BACKGROUND OF THE PRESENT INVENTION

DTXT Headgear such as hats, caps and bandannas are necessary in ...

DTXT The present invention joins together a conventional visor ...

DTXT U.S. Pat. No. 2,859,448 is for "Article of Headwear" issued...

... ..

DTXT U.S. Pat. No. 7,836,522 is for "Bandana Headwear with Flexible...

HDR1 SUMMARY OF THE PRESENT INVENTION

DTXT The present invention is an item of headgear comprised of a ...

HDR1 FIGURES OF THE PRESENT INVENTION

DTXT FIG. 1 is a flowchart showing the initial steps of constructing...

DTXT FIG. 2 illustrates step (180) of the flowchart shown in...

DTXT FIG. 3 illustrates step (190) of the flowchart shown in...

... ..

DTXT FIG. 8 illustrates step (330) of the flowchart shown...

HDR1 DETAILED DESCRIPTION OF THE PRESENT INVENTION

DTXT The present invention is a visored cloth headgear. A ...

DTXT The materials and components mentioned in the figures below ...

DTXT FIG. 1 is a flowchart showing the initial steps of constructing...

DTXT FIG. 2 illustrates step (180) of the flowchart shown in ...

DTXT FIG. 3 illustrates step (190) of the flowchart shown in...

... ..

DTXT FIG. 8 illustrates step (330) of the flowchart shown in...

DTXT Dimensions of the pieces that make up the present invention may...
DTXT It should be understood that the present invention is a visored...

From MAPS-XML file for **PATB US8621667B2**:

<desc>

```

<dtxt>The present invention is an item of headgear comprised of ... hat.</dtxt>
<hdr1>BACKGROUND OF THE PRESENT INVENTION</hdr1>
<dtxt>Headgear such as hats, caps and bandanas are necessary in ... sun.</dtxt>
<dtxt>The present invention joins together a conventional visor ... hat.</dtxt>
<dtxt>U.S. Pat. No. 2,859,448 is for &quot;Article of Headwear&quot;... visor.</dtxt>
<dtxt>U.S. Pat. No. 5,542,127 is for &quot;Combined Bandana and ... itself.</dtxt>
<dtxt>U.S. Design Pat. No. Des. 419,750 is for a &quot;Combined ... visor.</dtxt>
<dtxt>U.S. Pat. No. 7,836,522 is for &quot;Bandana Headwear with ... visor.</dtxt>
<hdr1>SUMMARY OF THE PRESENT INVENTION</hdr1>
<dtxt>The present invention is an item of headgear comprised of a ... hat.</dtxt>
<hdr1>FIGURES OF THE PRESENT INVENTION</hdr1>
<dtxt>FIG. 1 is a flowchart showing the initial steps of ... invention.</dtxt>
<dtxt>FIG. 2 illustrates step (<b>180</b>) of the flowchart shown in FIG. 1.</dtxt>
<dtxt>FIG. 3 illustrates step (<b>190</b>) of the flowchart shown in FIG. 1.</dtxt>
...
<dtxt>FIG. 8 illustrates step (<b>330</b>) of the flowchart ... (<b>98</b>).</dtxt>
<hdr1>DETAILED DESCRIPTION OF THE PRESENT INVENTION</hdr1>
<dtxt>The present invention is a visored cloth headgear. A ... hat.</dtxt>
<dtxt>The materials and components mentioned in the figures ... (<b>30</b>).</dtxt>
<dtxt>FIG. 1 is a flowchart showing the initial steps of ... (<b>100</b>).</dtxt>
<dtxt>FIG. 2 illustrates step (<b>180</b>) of the flowchart ... shown.</dtxt>
<dtxt>FIG. 3 illustrates step (<b>190</b>) of the flowchart shown in ... </dtxt>
...
<dtxt>FIG. 8 illustrates step (<b>330</b>) of the flowchart shown ... .</dtxt>
<dtxt>Dimensions of the pieces that make up the present invention ... .</dtxt>
<dtxt>It should be understood that the present invention is a ... folds.</dtxt>

```

</desc>

Claims Text Group

The Claims Text Group consists of the text of the numbered claims and additional claims elements when a numbered claim is divided into multiple text phrases. Each numbered claim is technically a sentence that ends with a period. Periods in a claim other than the final period should only be for abbreviations, the decimal point in a real number or another legal use other than the end of a sentence. A detailed discussion on claims text formatting can be found above in the Claims Text section on page 28.

Claims Section

| | | | | |
|-------------|------------|------------|-----------------------------------------------------------|----------------------------------------------------------------------------------------------|
| CLMS | ALL | STL | Claims Example: CLMS We hereby Claim: | A section label followed by 0 or more statements such as "We Claim" - "I claim", etc. |
|-------------|------------|------------|-----------------------------------------------------------|----------------------------------------------------------------------------------------------|

| | | | | | |
|-------------|--------------------|-----|--|----------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CLMF | EPA EPB | STL | | Claims In the FRENCH language* | Claims text in French * Claims text sections are included in EP Patents in 3 languages (en, de, and fr)! |
| CLMD | EPA EPB | STL | | Claims In the GERMAN language* | Claims text in German * Multiple Description sections are allowed in EPs in different languages! |
| CNTX | ALL | RT | | Numbered Claims text | Paragraph (1 to n claims) |
| CTXT | ALL | RT | | CTXT another text element of a numbered claims; | Claims Text Element (part of previous numbered claim text paragraph) |
| DCLM | USB | S | | DCLM No claims statement for Designs. | Begin Design Claim Section - No statement such as "I claim..." exists in a design claim section |
| DCTX | USB | RT | | DCTX Ornamental design for a sneaker as shown in the accompanying drawings. | Single, Unnumbered Claim – in a Design patent. |
| TABL | ALL | RT | | TABL Spaces are significant (keep all space characters) | Table Text Line – Tables are created as added Claim Text Elements similar to CTXT . |
| HDR1 | ALL | RT | | HDR1 Heading Text (Level 1) | Heading Text Line entry HDR2 through HDR6 are defined <u>but will not be used in the Claims Section</u> . Claims and Amended Claims will only use HDR1 . |
| CMNT | ALL | RT | | CMNT Comments are not part of the specification and are occasionally inserted to provide additional information for the user processing the file. | Comment Line Editorial comments added that are not part of the specification. |

Example of a typical Claims Section with long lines truncated (...), and a few lines in each area removed (... ..) copied from MAPS file for **PATB US8621667B2**:

```

...
CLMS I claim:
CNTX 1. A visored cloth headgear, comprising:
CTXT a visor bill cover;
CTXT a visor bill, said visor bill permanently disposed inside said...
CTXT a cloth, said cloth attached to said visor bill cover;
CTXT a tie, said tie formed from said cloth;
...
CTXT wherein said visor bill and said visor bill cover are fixed in...
CTXT wherein said cloth is only bound to said visor bill cover, providing ...
CNTX 2. The visored cloth headgear of claim 1, further comprising a top ...
CNTX 3. The visored cloth headgear of claim 2, wherein said top lip and ...
CNTX 4. The visored cloth headgear of claim 1, further comprising a slit...
...
CNTX 8. The visored cloth headgear of claim 1, wherein said cloth has ...
CNTX 9. A visored cloth headgear, comprising:
CTXT a visor bill cover;
CTXT a visor bill, said visor bill permanently disposed inside said ...
CTXT a cloth, said cloth attached to said visor bill cover;

```

CTXT a tie, said tie formed from said cloth;
 ...
 CTXT wherein said visor bill and said visor bill cover are fixed ...
 CTXT wherein said cloth is only bound to said visor bill cover, ...
 PATE

From MAPS-XML file for PATB **US8621667B2**:

```
<clms>
  <stmt>I claim:</stmt>
  <cntx>1. A visored cloth headgear, comprising:</cntx>
  <ctxt>a visor bill cover;</ctxt>
  <ctxt>a visor bill, said visor bill permanently disposed inside said ... </ctxt>
  <ctxt>a cloth, said cloth attached to said visor bill cover;</ctxt>
  <ctxt>a tie, said tie formed from said cloth;</ctxt>
  ...
  <ctxt>wherein said visor bill and said visor bill cover are fixed in ... </ctxt>
  <ctxt>wherein said cloth is only bound to said visor bill cover, ... </ctxt>
  <cntx>2. The visored cloth headgear of claim 1, further comprising a top ... </cntx>
  <cntx>3. The visored cloth headgear of claim 2, wherein said top lip and ... </cntx>
  <cntx>4. The visored cloth headgear of claim 1, further comprising a slit ...</cntx>
  ...
  <cntx>8. The visored cloth headgear of claim 1, wherein said cloth has ... </cntx>
  <cntx>9. A visored cloth headgear, comprising:</cntx>
  <ctxt>a visor bill cover;</ctxt>
  <ctxt>a visor bill, said visor bill permanently disposed inside said ... </ctxt>
  <ctxt>a cloth, said cloth attached to said visor bill cover;</ctxt>
  <ctxt>a tie, said tie formed from said cloth;</ctxt>
  ...
  <ctxt>wherein said visor bill and said visor bill cover are fixed ... </ctxt>
  <ctxt>wherein said cloth is only bound to said visor bill cover, ... </ctxt>
</clms>
```

From MAPS Design Patent **USD697067S1**:
 (the last 2 lines of the DESC section were left intact for context)

...
 DTXT FIG. 15 is a left side view of the portable terminal of FIG. 9; and,
 DTXT FIG. 16 is a right side view of the portable terminal of FIG. 9.
 DTXT The portions shown in broken line are not claimed.
 DCLM
 DCTX CLAIM We claim the ornamental design for a portable terminal, as shown
 and described.
 PATE

Note: The MAPS code **PATE** (Patent End) was left intact in the above examples for context.

From MAPS-XML Design Patent **USD697067S1**:

```
<dtxt>FIG. 16 is a right side view of the portable terminal of FIG. 9.</dtxt>
<dtxt>The portions shown in broken line are not claimed.</dtxt>
</desc>
<clms>
```

```

    <dcxt>CLAIM We claim the ornamental design for a portable terminal, as shown
        and described.</dcxt>
</clms>
</pub-us-grant>
</ipdata-maps-xml>

```

NOTES:

1) The MAPS-XML Design Claim Section uses **<clms>** which is the same Group Tag used for claims in all English Applications and Granted patents. The DCLM MAPS tag was a hold-over from APS to differentiate the single unnumbered Design claim from numbered Utility patent claims. **<dctx>** was kept for the claims text itself, since the single Design claim is NOT a numbered claim.

Search Report Group

| Search Report Section | | | | | |
|--------------------------------|----------|----|--|----------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| SRPT | EP WO | S | | SRPT Begin International Search Report | Begin Search Report Section Included (usually OCR) in some PCT applications of Kind Code A1 and A3, and some A1 and A3 EP pubs. |
| STXT | EP WO | RT | | STXT a line or paragraph of text in a Search Report | This is a text paragraph in any language contained in the Search Report. |
| TABL | ALL | RT | | TABL Spaces are significant | Table Text Line – Tables are created as added Claim Text Elements similar to CTXT . |
| HDR1 | ALL | RT | | HDR1 Heading Text (Level 1) | Heading Text Line entry HDR2 through HDR6 are defined but will not be used in the Claims Section . Claims and Amended Claims will only use HDR1. |
| CMNT | ALL | RT | | CMNT Comments are not part of the Claims Text specification and are usually inserted to provide additional information for the user processing the file. | Comment Line |
| Family Document Section | | | | | |

Search Report Section

| | | | | | |
|------|-----|----|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| DFAM | ALL | RD | | <p>DFAM US8117319B2 20120214</p> <p>The International Search Report data lists family documents, and this MAPS code will be used for these publication numbers when we redesign the Search Report sections to hold the various elements. Search Report text is still OCR sourced, but the EPO is moving towards electronic copies of the data.</p> <p>The WIPO PDF OCR DVD subscription contains all of the search report data, but the quality is so poor it is not yet worth pursuing the data - 1 of every 3 pub numbers is incorrect (<i>i.e.</i>, - damaged from bad OCR)</p> | <p>Document In Family</p> <p>One document number per line. The number may be followed by a space and the publication date, format YYYYMMDD.</p> |
|------|-----|----|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|

See **Appendix S** – titled **Search Report Text Data Format**

General Text Codes Used in One or More Sections

| | | | | | | | | | | | | | | | | | |
|----------------------------------------------|----------|-----------|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|-----------|--------|----|-----|--------|----|-----|--------|----|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| TABL | ALL | RT | | <p>Example:</p> <p>TABL Table US-00001</p> <table border="1"> <tr> <td>TABL Item</td> <td>Water(%)</td> <td>Sodium(%)</td> </tr> <tr> <td>TABL 1</td> <td>29</td> <td>2.5</td> </tr> <tr> <td>TABL 2</td> <td>35</td> <td>1.4</td> </tr> <tr> <td>TABL 3</td> <td>19</td> <td>3.5</td> </tr> </table> | TABL Item | Water(%) | Sodium(%) | TABL 1 | 29 | 2.5 | TABL 2 | 35 | 1.4 | TABL 3 | 19 | 3.5 | <p>Indicates One line of table data.</p> <p>All items are columnar, all spaces significant. Best displayed with a mono-spaced font.</p> <p>Tables may appear in DESC, CLMS, ABST and SRPT sections.</p> |
| TABL Item | Water(%) | Sodium(%) | | | | | | | | | | | | | | | |
| TABL 1 | 29 | 2.5 | | | | | | | | | | | | | | | |
| TABL 2 | 35 | 1.4 | | | | | | | | | | | | | | | |
| TABL 3 | 19 | 3.5 | | | | | | | | | | | | | | | |
| HDR1 HDR2 HDR3 HDR4 HDR5 HDR6 | ALL | RL | | <p>HDR1 Background</p> <p>HDR1 Summary</p> <p>HDR1 Detailed Description</p> <p>Six different Heading levels are Defined in MAPS V7.0a. Only HDR1 is current in use. This will change over 2016 as we upgrade software systems.</p> | <p>Text Heading used to define the following text sections or subsection of text.</p> <p>Usually used as a header line in a table – Label lines are not indexed. Usually a header line or a divider such as dashes in a table or similar (not used in a search – not generally indexed)</p> | | | | | | | | | | | | |

| | | | | |
|------|-----|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CMNT | ALL | RT | <p>MAPS Comment Field This is a comment field that may contain data not included as part of the original document. It usually indicates a comment made by an editor (after publication), or if a section of a patent document indicates a section is not part of the official patent or application, such as this disclaimer at the very end of an EP patents or Application: CMNT This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document.</p> | <p>The CMNT can also be used to note the source of the document, errors in OCR text, Original document numbers, formatting issues, etc.. Used for EP000 Office Management Fields and also in the EP Reference Lists that are not part of the EP publications (i.e., provided for convenience).</p> |
|------|-----|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

End of Patent

| | | | | |
|------|-----|---|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|
| PATE | ALL | S | <p>PATE End of Patent Document US printed patents always had the 5 double spaced asterisks centered to herald the end of the text: * * * * * This ensured the reader that the last page as intact. That is also the the purpose of the MAPS Code PATE</p> | <p>Patent End. Used to confirm that the file has not been prematurely cut off.</p> |
|------|-----|---|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|

Miscellaneous Comment Entries (CMNT)

The following **CMNT entries** (added in MAPS 6.3 and continued in MAPS 7.0a) will appear in US MAPS Applications when the corresponding Data Flags are present in the USPTO XML data. The **US Publication Filing Type** and **Rule 47** text are **not** printed on printed published applications and are provided to MAPS users for information purposes only. The **US Publication Filing Type** is used to determine document **Kind Code** and possibly other document status indications. A comment indicates that the “Redacted” type means that some application information was intentionally left out of the published application (it does not indicate what information it might be).

One of the following 6 Comments may appear in the MAPS data for published applications near the top of the Bib section:

CMNT US Publication Filing Type: us-original-publication-voluntary

CMNT US Publication Filing Type: us-original-publication-redacted

CMNT US Publication Filing Type: us-original-publication-amended

CMNT US Publication Filing Type: us-republication-corrected

CMNT US Publication Filing Type: us-republication-redacted

CMNT US Publication Filing Type: us-republication-amended

CMNT This application was filed under Rule 47, indicating the applicant(s) refused to execute the application or could not be found.

The following **CPA text** will appear as a comment on the face of applicable published applications:

CMNT This is a publication of a continued prosecution application (CPA) filed under 37 CFR 1.53(d).

It will also be present in the MAPS files for that document.

End of MAPS CODE Table

APPENDIX A – Name and Address Data for Concerned Parties

Background Information

When MAPS was first designed, the USPTO was providing complete party details including **last name, first name, address, city, state, and country**. Our initial source for EP and PCT data was basically Last Name and Initials, plus the Country. Our second set of data in 2005 and 2006 added the first name and the City. Then in 2009/2010, the EPO rebuilt all of their data, and we finally had complete address data for the parties, and even more important than that, we had decent full text. A year later, after WIPO had taken over their own data maintenance and distribution, they rebuilt their entire text (albeit OCR) and not nearly as nice as the EPO or USPTO data. However, as the EPO and WIPO added more details with addresses, the USPTO removed all address data and provide only the Name, City and State (US) or Country Non-US.

Right after the EPO and WIPO text-upgrades, we needed a place to put all of the new, clean, data, and MAPS 6.0 was the answer. We are now at **MAPS 7.0a** and we have a new MAPS-XML product available, and overall, the data quality has greatly improved. The greatly improved party data drove the creation of the extended MAPS Codes for the Party Data that appeared in MAPS 6.0, and are described below.

MAPS Party Fields with Name and Address Data

As the various patent authorities continue to add, delete or change items in the Party data, we have stabilized our design and we include as much of the party data as possible (as they provide) into our *Extended* format. The four MAPS codes in our extended format that document the APPLICANT, INVENTOR, AGENT and ASSIGNEE are:

- **INVX** Inventor (with more defined detail)
- **APLX** Applicant (with more defined detail)
- **AGTX** Agent (with more defined detail)
- **ASGX** Assignee or Grantee (with even more detail)

How the Party Fields are Divided

The four main Parties all have the five sub-sections divided by four vertical bars shown in the Inventor Example (**INVX**) below:

INVX Name | Address | City, State or Province, Region | Country | Postal-Code

Name (Person or Entity/Organization)

Most Authorities break out the Last and First name (and often the middle initial), but some have a generic “name” field (JP). They all have the Last Name first and it is usually separated from the first name by a comma, but some data may contain a semicolon. The **APLX** and **ASGX** may be an individual or an organization (company, corporation, university, etc.). The **AGTX** (agent) may be an individual (attorney), a law firm or similar.

Address

Address data is no longer provided by the USPTO (discontinued in 2011) and is also not included in the JP abstract data. The EPO and WIPO still include a fairly complete address although the data in the WO files is often in the wrong fields. We use several custom heuristic algorithms to try to get the misplaced data where it belongs, and after two years of fine-tuning, they work fairly well.

City, State and Region

City and State, or City and Region, where applicable, is included for all parties from all Authorities except older JP Abstract data. There will always be a comma between a US city and the state, and between a city and region (e.g., Vancouver, BC). U.S. states may be spelled out or they may be two letters.

Country per ST.3 Standard

The country will always be two capital letters per WIPO Standard ST.3. If the country is not supplied, two

dashes (|--|) will replace the two capital letters.

Postal-Code

Most countries have postal codes that are a fixed format and we have algorithms that check and formats (or reformat) the codes for the 17 most listed countries.

Additional Elements in (APLX) Applicant (all authorities)

The Applicant (**APLX**) also has nationality and residence fields which are two letter country codes separated by a comma in square brackets that immediately follow the postal-code. The Nationality has been removed from US data and the Residence has been made "Optional" but is still included with most US applicants. Two Dashes (2D hex) will be inserted if the country is not supplied. PCT and EP applicants still include both elements (Nat & Res). Example (with non-germane data removed):

```
APLX ... |||| [US,US]
```

US and PCT Applicants (**APLX**) may also contain up to 3 additional elements **after the [nat, res] field in parenthesis** that further describes the Applicant's status. This information does not appear on the face of the publication, but may be implied by way of a different INID code in front of the Applicant name (see 71 to 75 in the INID table earlier in this specification). Examples:

```
APLX ... |||| [US,US] (app-Type: applicant-inventor, applies to states: all)
APLX ... |||| [US,US] (app-Type: applicant, Applies to states: all-except-us)
APLX ... |||| [US,US] (app-Type: applicant-inventor, applies to states: us-only)
```

Applicants on US applications may also contain an additional category describing the authority of the applicant if it is not the inventor. The element is included after another comma, and is named **applicant-authority-category**: and will be one of the following:
inventor, legal-representative, party-of-interest, obligated-assignee or assignee.

An example of a US Application with all three elements (line feeds added for clarity):

```
APLX ... |||| [US,US] (app-Type: applicant-inventor,
                      applies to states: us-only,
                      applicant-authority-category: assignee)
```

Additional Elements in EP (INVX) Inventor

EP Granted patents may contain a field that lists states (countries) where a transfer of rights has taken place regarding the patent. An optional date, which is the **Effective Date of Rights**, may also be included, but usually isn't. If not included, the effective date should be the date of the Grant. However, please check with your registered Agent to be certain. The States or extended States are listed in the following format:

```
[Desig. States: CC CC CC CC CC] Eff. Date of Rights: YYYYMMDD
[Desig. Ext. States: CC CC CC CC CC] Eff. Date of Rights: YYYYMMDD
```

Below are examples (data not germane to the example is shortened and replaced with ellipsis ...):

```
ASGX Unilever PLC [IID: 101...][IRF: ...]|Unilever ...|London EC4Y 0DY|| [Desig. States: CY GB IE]
ASGX Unilever NV. [IID: 100...][IRF: ...|Weena 455|3013 Rot...|NL| (cont. next line)
                      [Desig. States: AT BE BG CH CZ DE DK EE ES FI FR GR IT LI LU MC NL PT SE SK TR]
```

The Designated States are from EP field **B736EP**
The Designated Extended States are from EP field **B737EP**
The Optional Effective Date of Rights for both of the above is captured from EP field **B738EP**

Additional Elements in U.S. (ASGX) Assignee

In place of the the [Nat, Res] elements, the Assignee data may contain a two digit number in square brackets. This number describes what type of Assignee is it. The value is selected from the following list:

- 01** Unassigned;
- 02** United States company or corporation;
- 03** Foreign company or corporation;
- 04** United States individual;
- 05** Foreign individual;
- 06** U.S. Federal government;
- 07** Foreign government;
- 08** U.S. county government;
- 09** U.S. state government

Example of the **ROLE** in an Assignee entry:

ASGX Asahi Engineering Co., Ltd. |Ishikawa|JP| [03]

Additional Elements in (APLX, AGTX and ASGX) Applicant, Agent, Assignee

The Applicant (**APLX**), Agent (**AGTX**) and Assignee (**ASGX**) may also have one or two additional elements in square brackets at the end of the Name sub-section. The elements are the Individual-ID (**IID** or **iid**) and Individual Reference number (**irf** or **IRF**). For example:

APLX Panasonic Corp. [iid: 101389465] [irf: BB72366EP]|1006 Oaza Kadoma|Kadoma-shi, Osaka|JP|571-8501

The (**iid**) numbers are assigned by the EPO to Individuals or Entities in the various processes (application, Opposition, Assignment, etc.) to uniquely identify them. We assume it is a permanent assignment to the entity, but they don't specifically say (your Registered Agent will know).

The (**irf**) is assigned to a particular item such as a *Filing, an Opposition* or even an external (Non-Patent Reference) cited by an individual to ensure it can be located or referenced with no ambiguity in other processes. Details beyond what we have already described seem to be a closely guarded *State Secret* of the EU and the EPO.

The USPTO has recently added the **IID** field, but has not yet populated it.

Examples:

An Inventor from a PCT application:

INVX Smith, Robert D. |460 Cherokee Ln, Suite D|Tuscon, AZ|US|85702 [US, US] <EOL>

An Applicant from a PCT application:

APLX KIYOTA, Yukinori |c/o SONY Corp. 1-7-1 Konan Minato-ku|Tokyo|JP|108-0075<EOL>

An Assignee from a US Patent:

ASGX International Business Machines Corporation||Armonk, NY|US| [02] <EOL>

An Agent from and an EP Patent:

AGTX Boulton Wade Tennant [IID: 101370347]|Verulam Gardens, 70 Gray's Inn Rd|London|GB|WC1X 8BT

Additional Notes on Extended Party Fields:

1. Empty Fields – The 4 Vertical Bars will always be in the data line even when parts of the data are not populated. Vertical Bars found in the data will be converted to semicolons. For example: If we are only provided the name, city, state and country, the line will look like this:

INVX Smith, Bob D. | |Tuscon, AZ|US| [--,US]

3. Commas – Street Address entries may be composed of two and rarely three elements from the source XML data. We will place a comma between the lines of data within a single sub-field.

4. Country Codes – All two letter country entries will be listed in WIPO Standard ST.3. Older data may list country codes that no longer exist or have been changed (e.g., **SU** was the Soviet Union, UK was changed to **GB**, etc.).

5. Sir or last name – The Last or Sir name is usually first in the source data, and we will keep it that way. The last name will be separated with a semicolon or comma.

6. Postal Codes – We attempt to move the Postal Code into the correct sub-section when it is found elsewhere. But this is more of an art than a science since human beings are involved in the data entry. We have been working to perfect our validation and identification algorithms, and they are vastly improved, but are certainly not perfect by any stretch of the imagination.

7. Newer US Data – The very newest US Patent data usually **does not** have the street address or postal codes. While the EPO and WIPO continue to add and clean up their address data, the US is reducing the Party data as much as possible. Go figure...

8. Nationality and Residence or Role – When the two letter country code for either Nationality or Residence is replaced by “Omitted” or is not included at all in the source data, we will insert two dashes (UTF-8 single byte code 2D hex). For example: [--,US]. The US Application DTD Specs contained this note: “**Removed nationality element from us-applicant**” The USPTO hasn’t supplied the Nationality data for over 2 years now. So this is no surprise. But the EPO and WIPO still supply all or part of the data. The EPO is most consistent with both Nat and Res in their applicants, so we will keep the format as-is, for now.

9. Regions - The City-State sub-section occasionally has three elements in it. This happens when a country has a City, a Province and a Region,

10. Spaces after Vertical Bar separators – Some of our algorithms may add an additional space after the Vertical Bars that separate sub-fields, but this is not guaranteed. Empty sub-fields may contain a single space or they may not. Do not expect one, but be prepared if one is there. We do attempt to make MAPS lines legible so they can be edited by our own people to correct bogus OCR data (such as the data supplied by WIPO).

APPENDIX B – Country and Language Codes

Country and Language Codes listed in MAPS patent documents

The MAPS data currently covers PCT (WO), EP and US Applications, WO, EP, US and JP Abstracts, and EP and US granted patents.

Country Codes

Because of the maintenance involved in keeping the MAPS specification up to date, we refer you to the **WIPO ST.3** Recommended Standard which is the current list of **two-letter country codes** used in all patent publications around the world. All Standards including ST.3 can found found here:

http://www.wipo.int/standards/en/part_03_standards.html

Language Codes

Languages indicated in patent documents used the **two letter lower-case codes** from ISO-639, and below is a subset that should cover every language we run across in all MAPS data.

ISO 639 Codes (Names of Languages)

Technical contents of ISO 639:1988 (Subset) - "Code for the representation of names of languages".

| | | | | |
|-----------------|-----------------|----------------|-------------------|---------------|
| aa Afar | eu Basque | ka Georgian | oc Occitan | sv Swedish |
| ab Abkhazian | | kk Kazakh | om (Afan) Oromo | sw Swahili |
| af Afrikaans | fa Persian | kl Greenlandic | or Oriya | ta Tamil |
| am Amharic | fi Finnish | km Cambodian | | te Tegulu |
| ar Arabic | fj Fiji | kn Kannada | pa Punjabi | tg Tajik |
| as Assamese | fo Faeroese | ko Korean | pl Polish | th Thai |
| ay Aymara | fr French | ks Kashmiri | ps Pashto, Pushto | ti Tigrinya |
| az Azerbaijani | fy Frisian | ku Kurdish | pt Portuguese | tk Turkmen |
| ba Bashkir | | ky Kirghiz | | tl Tagalog |
| be Byelorussian | ga Irish | | qu Quechua | tn Setswana |
| bg Bulgarian | gd Scots Gaelic | la Latin | | to Tonga |
| bh Bihari | gl Galician | ln Lingala | rm Rhaeto-Romance | tr Turkish |
| bi Bislama | gn Guarani | lo Laothian | rn Kirundi | ts Tsonga |
| bn Bengali, | gu Gujarati | lt Lithuanian | ro Romanian | tt Tatar |
| Bangla | ha Hausa | lv Latvian, | ru Russian | tw Twi |
| bo Tibetan | hi Hindi | Lettish | rw Kinyarwanda | |
| br Breton | hr Croatian | mg Malagasy | sa Sanskrit | uk Ukrainian |
| | hu Hungarian | mi Maori | sd Sindhi | ur Urdu |
| ca Catalan | hy Armenian | mk Macedonian | sg Sangro | uz Uzbek |
| co Corsican | | ml Malayalam | sh Serbo-Croatian | vi Vietnamese |
| cs Czech | ia Interlingua | mn Mongolian | si Singhalese | vo Volapuk |
| cy Welsh | ie Interlingue | mo Moldavian | sk Slovak | |
| | ik Inupiak | mr Marathi | sl Slovenian | wo Wolof |
| da Danish | in Indonesian | ms Malay | sm Samoan | |
| de German | is Icelandic | mt Maltese | sn Shona | xh Xhosa |
| dz Bhutani | it talian | my Burmese | so Somali | yo Yoruba |
| el Greek | iw Hebrew | | sq Albanian | |
| en English | | na Nauru | sr Serbian | zh Chinese |
| eo Esperanto | ja Japanese | ne Nepali | ss Siswati | zu Zulu |
| es Spanish | ji Yiddish | nl Dutch | st Sesotho | |
| et Estonian | jw Javanese | no Norwegian | su Sundanese | |

APPENDIX C – Correction Codes for PCT (WO) Pubs

PCT Correction Codes used in ALL MAPS WO Pubs

The newer correction codes, that began to appear on the scene circa 2003, are now used in all PCT Correction publications. Older Correction code are no longer used in any MAPS publications. The tables and all references to them have been removed from this version of the Specification. **If you need the older correction codes for historical reasons please save copies of your older MAPS specifications.**

The current **2 part correction codes** have a single digit Correction Type, and a two character Republication Code (one letter and one digit). Most of these **Republication Codes** were standardized in 2001 and 2003, but it was 2009 before they were in full use, and 2011 before for the rest of the back files were completely updated, including all bibliographic and full text references.

Below is an example of a **CORR** entry in a MAPS file (typical A8 Kind publication):

```
CORR 4, R6, pages 1/29-29/29, drawings, replaced by new pages 1/29-29/29; due to late transmittal by the receiving Office
```

Here is another example with **two corrections listed in one document**. The number of CORR entries should equal the total number of corrections for the publication in which they are contained (including the current correction):

```
PATW WO2009053490A9
PBDT 20090430 <== Date of Original Publication
PKCD A9
MVER 70a-20150909
CORR 2, R3, Bulletin: 10/2011
CODT 20110310 <== Publication Date of Correction
CORR 3, R4, Bulletin: 39/2009
DSRP 20090924
APNO PCT/EP2008/064530
APDT 20081027
```

The Tables on this and the next page lists the TYPE and REPUBLICATION Codes

Table of Correction Types (single digit, 1 to 4)

| Value | Description |
|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Not a “correction,” but instead is a legal modification allowed under the PCT. The only use we’ve seen is the Amended Claims filed under PCT Article 19, but it may apply to others. |
| 2 | Complete replacement , but of a “further publication” such as an ISR |
| 3 | New, additional document added to the existing series for this publication. Often called a “ <i>further publication</i> ” such as a Kind A3 Search Report (ISR). |
| 4 | Replacement of a Complete Document or Front Page of Primary Publication (A1 and A2). Published as Kind A8 (front page only) or Kind A9 (complete document) publication. |

Table of Republication Codes (R1 to R9)

This table is the “cleanest” description of these codes we can find, and due to its age (2006) we have reviewed the PCT Correction Circulars to make sure it is accurate:

| Code | No. published through 2005 | Description of Current Republication Codes |
|-----------|----------------------------|---------------------------------------------------------------------------------------------------|
| R1 | 2,876 | Amended Claims – Publication of Art. 19 Amendment (only the claims are published) |
| R2 | 17 | Revised International Search Report (A1 or A3 Kinds) |
| R3 | 1,147 | Revised International Search Report with updated Bibliographic data (A1, A2 or A3 Kinds) |
| R4 | 126,529 | Initial publication of Search Report (A3 Kind for Prior A2 Kind) |
| R5 | 7,288 | Correction of Entry in Sect. 1 (Bib) |
| R6 | 11,435 | Full Reprint of publication |
| R7 | 2,104 | Corrected Version of a Front Page (not used after 2007 per PCT Circular 1127-76) |
| R8 | 1,535 | Later Publication Of A Declaration Under Article 17(2)(A) |
| R9 | 22 | Revised Version of Search Report Language Translation |
| | | |
| | | We left the <i>number published</i> column in to give you an idea of the popularity of each code. |

In (partial) compliance with WIPO Standard ST.16 [[PDF](#)], the following publication **Kind Codes** are used

- A1** Primary Level Publication with International Search Report
- A2** Primary Level Publication **without** International Search Report
- A3** Secondary or Further Publication with International Search Report
- A4** Publication of Amended Claims, 2009 and later
- B1** Publication of Amended Claims, prior to 2009
- A8** Front Page or Bibliographic Correction. New Front Page Published
- A9** Correction with Complete Replacement of Primary or Secondary Level Publication

Correction Codes are found in the A3, A4, B1, A8 and A9 PCT publications.

APPENDIX D – Dot Codes in Older US Data

Introduction

The now infamous USPTO “Green Book” (**GB**) is a 3 ring bound volume of approximately 900 pages that describes the APS patent text storage and display format designed by the USPTO.

APS is an acronym for *Automated Patent System* and was a computerized patent text storage system created by the USPTO in the mid 1970s. An important part for accurately storing patent data was Appendix C covering the Dot Code system. Appendix C in the Green Book described how special characters and symbols, along with certain text styles such as subscript and superscript, were represented in the non graphical environment of the mainframe computers and plain text based display systems of that era (circa 1975).

Using the limited number of codes in the EBCDIC system, the PTO needed a way to allow storage and searching of many special characters, symbols and styles that are used in patents, but not found in the EBCDIC or ASCII character tables. Appendix-C of the GB is how this was accomplished.

The methods and codes described in Appendix-C are still in use today by the USPTO on their web site and in all of the pre-year 2000 text that they distribute in bulk form. Even patent text files that were converted to other storage formats usually contained the codes described in Appendix-C because there was no standardized way to represent these codes and styles until SGML and XML came along, plus, it took a few years for those standards to settle out to the point they were dependable (a decent guarantee that entity code used will look the same a few years later, and that you won't find two or three different symbols using the same entity code).

Background

APS was an 80 column text format specifically designed to store and format patent text for display on the various computer systems of that era (early 1970s). The character codes used on the mainframes at the USPTO were EBCDIC using IBM “code tables” on IBM, and IBM compatible mainframes such as the *Amdahl* computers at the USPTO. Everything was processed in 80 column (or 80 byte) logical sections. Rudimentary text formatting was maintained through the use of codes that represented certain line indentations (e.g., 3 4 or 5 spaces preceding the indented text), spacing characteristics and line endings. The actual APS codes were 4 and 5 character codes that were found at the beginning of every physical line. Some were used to delineated logical sections of the patent, and others signified groups of lines, line types or lines in tables.

Since much more advanced software and display systems are now available for parsing and displaying the text data, many of the APS codes are no longer needed, such as those for spacing and page formatting. There was also a large number of codes never used, but still defined in the Dot Code system (at least we can't seem find them with today's search capabilities). Some say the system was “over designed” but we will take them on in a debate any time in support of a system that lasted 40 years (**and worked, all of those 40 years**). Was it perfect? What software is? The original analysts and programmers did an excellent job with that system we heartily commend them on a job well done. Software that last 10 years borders on a miracle, these days.

Subset of Green-Book Appendix

In this appendix, we have include a subset of Dot Codes that we know are still in use in our MAPS files. As near as we can tell, only original copies of the GB printed by the US Government Printing Office had all of the the graphical images of the special symbols and characters located in column 1 of the main table (below). Most copies of the GB that have been distributed by the USPTO since about 1994 have been text copies run off on high speed line and laser printers, and in the first column where the symbols *should be printed* are found the “call out” codes used by GPO. *Call Out Codes* are numbers preceded by a plus sign (+) to indicate which symbol to print in its stead. IP Data Corporation has added the symbols in a new column (*column 2 in this document*) where we have been able to confirm their appearance by finding them in use in facsimiles of actual granted patents. When we could not find them in use, we left them blank or inserted ??? (3 question marks instead). This does not means they were not used, but simply means that we could not find one in all of the text provided by the USPTO. Some of the text we have was acquired by OCR and several different human editors verified the text. Symbols may have gotten by them (we did say they were “human,” right?).

Some notes on our table:

- 1) We added most of the symbols after locating them in a text search for the Dot Codes.
- 2) All EBCDIC Hexadecimal character codes in Column No. 4 were converted to ASCII Hex codes.
- 3) We added some of the editorial comments in a few code descriptions (column 5).

The “DOT CODES”

Having only a limited number of characters in the EBCDIC system, the USPTO engineers adopted a relatively common method in use in many systems back in the 1970s. This method simply uses a unique pattern of characters from the existing character set to identify new items that are needed but not available in the set.

The method that we call “Dot Codes” are words (strings of 1 or more characters) surrounded by periods (or “dots,” if you will) that name the unique symbol or character to be defined. The pattern itself (dot word dot with no spaces) **does not appear normally in properly formatted English text**. This makes it very easy to add all sorts of new symbols using the **Dot Code** pattern. For instance, the *degree* symbol (° - a small superscript circle usually following a numeric value) is used commonly in many patents, but it does not exist in the EBCDIC or ASCII character sets. This was defined as **.degree.** in the PTO Dot Codes (with no spaces between the word and the dots on either side of the word).

The USPTO could have demanded that every symbol and non-standard character on a patent application be spelled out long hand, but they knew they would soon find out that not everyone calls symbols by the same name, and in fact, there are multiple names for many math and scientific symbols (as you probably already know). However, a table, such as this Appendix-C, with a fixed list of the accepted symbols along with a corresponding unique DOT CODE leaves no room for ambiguity in the APS system, and that is why it was done this way.

This is not to say that there aren’t problems with this type of system. On the contrary, there were many (and some of them still haunt us today). Bear in mind that it was *humans performing the translations*, and some of them took some liberties such as seeing a code that was “close enough for government work” and assigning it when it wasn’t exactly the correct code to use, or simply making an honest mistake. However, we don’t think that the level of mistakes exceeded the normal, acceptable level of errors found in any normal document creation process.

The following table shows how special characters and symbols are represented in the text sections of APS and MAPS formatted text files. APS is an older text format for patent data that is still currently in use in various systems, and there is even unsupported software (including source code) available on the USPTO FTP site to translate data to and from various APS formats that were provided to the public over the years (<ftp://ftp.uspto.gov/pub/patdata/fms/>).

You will still see the Dot Codes in many systems today (in addition to the data that we provide). These include the USPTO web site search system, and the “East” and “West” search systems in the USPTO public search rooms. Just keep in mind: “*APS is dead. Long Live APS!*”

USING THE DOT CODES

The Dot Codes in this book are found all through the MAPS and APS patent text and where the codes have not been converted to UTF-8 display codes, they have not been modified. It is possible that spaces have been added before or after the text in patents after the year 1999 due do the fact that conversion software used may add a space in the process of turning 80 column line oriented text into variable length paragraphs. This should have little or no effect when parsing and indexing the MAPS text as these codes are fairly unique being surrounded by dots (periods) and quite often the code describes the symbol by name.

How you use the dot codes depends on how your system parses and indexes the Dot Codes and the special characters. Here are a few examples of Dot Codes taken from MAPS Patent text:

In the MAPS Text:

H.sub.2 0
f1 .apprch. f2
.pi. r .sup.2

In the Patent Facsimile:

H₂ 0 (subscript)
f₁ → f₂ (approaching)
π² (superscript)

Setting Up or Programming a Parser/Indexer for Dot Codes

How you set up your parsing system in order to index the Dot Code symbols will generally depend on two things:

- 1) What your indexing parser allows in the way of punctuation and characters, and
- 2) How you want to search for them in your system.

Some systems allow you to retain certain types of punctuation based on *Rules* set up in your lexical parser. You may want to set it up to keep the “Dots” on either side to force the unique identification of the symbols which may also allow you to program your display to insert the real symbols when building pages of text for display, or you may want to index them as standard words by discarding the dots, altogether. This makes searching easier, but may possibly cause confusion when the patent text is displayed. It all depends on your system, and if it isn’t the P9 system, you are on your own.

Generally, searching for things like formulas is not a standard practice in patent searching due to the many ways a single formula can be written, but you may find that you want to do it for some reason. If so, you need to know how the dot codes are handled in the search system you are using, and maybe even how they were parsed and indexed, as well.

On the USPTO Web Site

Even today (at least through 2014), if you look at the USPTO web site, their display code places spaces before and after most dot codes even if there were no spaces in the original APS text. This means that standard “text” searches should work fine on their web site, but some of the patents tend to look a bit confusing because of this. This is an anomaly caused by HTML and the way browsers handle spacing.

The DOT CODE SYMBOL/CHARACTER TABLE

Column 1 shows the “Call Out” number used for photo-composition when they printed the Green Book.

Column 2 (Symbol) shows the appearance of special characters on the patent Full-text/APS file.

Column 3 is the DOT code, if assigned one.

Column 4 is the Hex Code in ASCII (no DOT Code assigned to these)

Column 5 is the text description of the symbol

Notes:

- Missing numbers in column 1 (not in sequence) means the number is not used.
- ??? (3 question marks) in the Symbol column means we could not find a proper example of use
- Most Subscript and Superscript versions of symbols are not reproduced.
(See the original symbols for appearance).

DOT CODE TABLE

| Call Out# | Symbol (Image) | DOT CODE | ASCII Hex | Description & Comments |
|-----------|----------------|--------------------|-----------|-------------------------------------------------------|
| +1 | ` | | 60 | Single open quote |
| +2 | ' | | 27 | Single close quote |
| +3 | ' | | 27 | apostrophe |
| +4 | ~ | .varies. | | Varies as |
| +5 | ≡ | .ident. | | Identical |
| +6 | ≠ | .notident. | | Not identical |
| +7 | ∂ | .differential. | | Differential |
| +8 | [| | 5B | Open bracket |
| +9 |] | | 5D | Close bracket |
| +10 | Invisible | (width of EM Dash) | | Em space (wide space) |
| +11 | Invisible | (width of EN Dash) | | En space (narrow space) |
| +12 | Invisible | | | Thin space (even more narrow) |
| +13 | — | | | Em dash (width of two En dashes) |
| +14 | - | | | En dash |
| +15 | .. | .En. | | leader |
| +16 | ≠ | .noteq. | | Not equal symbol |
| +17 | ∫ | .integ. | | Integral symbol |
| +18 | ~ | .about. | | Roughly Similar (the symbol is what most entries had) |
| +19 | · | .multidot. | | Multiplication dot |
| +20 | ° | .degree. | | degree symbol |
| +21 | < | | 3C | Less than symbol |
| +22 | > | | 3E | Greater than symbol |
| +23 | ≤ | .ltoreq. | | Less than or equal to symbol |
| +24 | ≥ | .gtoreq. | | Greater than or equal to symbol |
| +25 | ≡ | .congruent. | | Congruent symbol |

| | | | | |
|-----|----------------------|---------------|----|----------------------------------------------|
| +26 | \cong | .apprxeq. | | Approx. equal symbol (also found \approx) |
| +27 | ¢ | .cent. | | Cent sign |
| +28 | ∞ | .infin. | | Infinity symbol |
| +29 | $\sqrt{\quad}$ | .sqrt. | | Square root symbol |
| +30 | + | | 2B | Plus sign |
| +31 | - | | 2D | Minus sign |
| +32 | = | | 3D | Equal sign |
| +33 | \times | .times. | | Multiplication sign |
| +34 | \div | .div. | | Division sign |
| +35 | \pm | .+-. | | Plus or minus sign |
| +36 | \mp | .-+. | | Minus or plus sign |
| +37 | " | | | Double open quote |
| +38 | " | | | Double close quote |
| +39 | \rightleftharpoons | .revreaction. | | Reversible reaction symbol |
| +40 | ' | .prime. | 7D | Prime symbol |
| +41 | " | .dblprime. | 7F | Double prime symbol |
| +42 | Δ | .increment. | | Increment symbol |
| +43 | \rightarrow | .fwdarw. | | Forward arrow (right) |
| +44 | \therefore | .thrfore. | | Therefore symbol |
| +45 | \because | .because. | | Because symbol |
| +46 | # | .music-sharp. | | Music sharp symbol |
| +47 | b | .music-flat. | | Music flat symbol |
| +48 | ? | | 3F | Question mark |
| +49 | ! | | 2I | Exclamation point |
| +50 | \equiv | .dbd. | | Double bond symbol (atomic/chemical) |
| +51 | | .vertline. | 7C | Vertical line symbol |
| +52 | \uparrow | .uparw. | | Up arrow |
| +53 | \sqsubset | .hoarfrost. | | Hoarfrost symbol |
| +54 | $\sqrt[3]{\quad}$ | .cuberoot. | | Cube root symbol |
| +55 | { | | 7B | Open brace |
| +56 | } | | 7D | Close brace |
| +57 | \equiv | .tbd. | | Triple bond symbol (atomic/chemical) |
| +58 | \square | .quadrature. | | Quadrature symbol |
| +59 | ∇ | .gradient. | | Gradient symbol (Nabla) |
| +60 | α | .alpha. | | Greek alpha symbol (lower - upper: A) |
| +61 | \oplus | .sym. | | Positive earth symbol (symmetry) |
| +62 | β | .beta. | | Greek beta symbol (lower - Upper: B) |
| +63 | \ominus | .crclbar. | | Negative earth sym. (circle bar) |
| +64 | Γ | .GAMMA. | | Greek GAMMA symbol (upper case) |
| +65 | γ | .gamma. | | Greek gamma symbol (lower case) |
| +66 | Δ | .DELTA. | | Greek DELTA symbol (upper case) |
| +67 | δ | .delta. | | Greek delta symbol (lower case) |
| +68 | ϵ | .epsilon. | | Greek epsilon sym. (lower case) |
| +69 | E | .EPSILON. | | Greek EPSILON sym. (upper case) |
| +70 | ζ | .zeta. | | Greek zeta symbol (lower case) |
| +71 | f | .function. | | Function symbol (or Function of) |
| +72 | η | .eta. | | Greek eta symbol (Upper: H) |
| +73 | Θ | .THETA. | | Greek THETA symbol (upper case) |
| +74 | θ | .theta. | | Greek theta symbol (lower case, |
| +75 | \downarrow | .dwnarw. | | Down arrow |
| +76 | ι | .iota. | | Greek iota symbol (Upper: I) |
| +77 | \leftarrow | .rarw. | | reverse arrow - left |
| +78 | κ | .kappa. | | Greek kappa symbol (Upper: K) |

| | | | |
|------|---|-----------------|--------------------------------------------|
| +79 | Λ | .LAMBDA. | Greek LAMBDA symbol (upper case) |
| +80 | λ | .lambda. | Greek lambda symbol (lower case) |
| +81 | | .parallel. | Parallel symbol |
| +82 | μ | .mu. | Greek mu symbol |
| +83 | | .sup..parallel. | Superscripted parallel symbol |
| +84 | ν | .nu. | Greek nu symbol (Upper: N) |
| +85 | Ξ | .XI. | Greek XI symbol (upper case) |
| +86 | ξ | .xi. | Greek xi symbol (lower case) |
| +87 | | .sub..parallel. | Subscripted parallel symbol |
| +88 | ο | .omicron. | Greek omicron symbol (lower case) |
| +89 | - | .sup..fwdarw. | Superscripted forward arrow |
| +90 | π | .pi. | Greek pi symbol (lower case) |
| +91 | - | .sub..fwdarw. | subscripted forward arrow |
| +92 | ρ | .rho. | Greek rho symbol (lower case) |
| +93 | Σ | .SIGMA. | Greek SIGMA symbol (upper case) |
| +94 | ς | .sigma. | Greek sigma symbol (lower case) |
| +95 | - | .sup..rarw. | Superscripted reverse arrow |
| +96 | τ | .tau. | Greek tau symbol (lower) |
| +97 | Υ | .UPSILON. | Greek UPSILON symbol (upper case) |
| +98 | υ | .upsilon. | Greek upsilon symbol (lower case J) |
| +99 | - | .sub..rarw. | Subscripted reverse arrow |
| +100 | φ | .phi. | Greek phi symbol (lower case) |
| +101 | Φ | .PHI. | Greek PHI symbol (upper case) |
| +102 | Χ | .chi. | Greek chi symbol (upper - lower undefined) |
| +104 | ψ | .psi. | Greek psi symbol (lower case) |
| +105 | Ψ | .PSI. | Greek PSI symbol (upper case) |
| +106 | Ω | .OMEGk. | Greek OMEGA symbol (upper case) |
| +107 | ω | .omega. | Greek omega symbol (lower case) |
| +108 | • | .cndot. | Solid dot - center |
| +109 | ⊞ | .andgate. | And gate symbol (logical AND: ∧) |
| +110 | ∠ | .angle. | Angle symbol |
| +111 | | .sup.(| |
| +112 | | .sup.] | |
| +113 | | .sub.(| NOTE: Column 1 numbers 111 thru 184 |
| +114 | | .sub.] | (except 129) represent special |
| +115 | | .sup.+ | characters as previously |
| +116 | | .sup.- | defined but appearing either |
| +117 | | .sub.+ | subscripted or superscripted. |
| +118 | | .sub.- | |
| +119 | | .sup.+- | |
| +120 | | .sup.-+ | |
| +121 | | .sub.+- | |
| +122 | | .sub.-+ | |
| +123 | | .sup.* | |
| +124 | | .sub.* | |
| +125 | | .sup.= | |
| +126 | | .sub.= | |
| +127 | | .sup.- | |
| +128 | | .sub.- | |
| +129 | ‰ | .permill. | Per mill symbol (salinity) |
| +130 | | .sup.1/8 | |
| +131 | | .sub.1/8 | |
| +132 | | .sup.3/8 | |

| | | | |
|------|---|-----------------|-----------------------------------|
| +133 | | .sub.3/8 | |
| +134 | | .sup.5/8 | |
| +135 | | .sub.5/8 | |
| +136 | | .sup.7/8 | |
| +137 | | .sub.7/8 | |
| +138 | | .sup.1/3 | |
| +139 | | .sub.1/3 | |
| +140 | | .sup.2/3 | |
| +141 | | .sub.2/3 | |
| +142 | | .sup.1/4 | |
| +143 | | .sub.1/4 | |
| +144 | | .sup.1/2 | |
| +145 | | .sub.1/2 | |
| +146 | | .sup.3/4 | |
| +147 | | .sub.3/4 | |
| +148 | | .sup..lmtoreq. | Superscripted less than or equal |
| +149 | | .sub..lmtoreq. | Subscripted less than or equal |
| +150 | | 1/8 | |
| +151 | | .sup..gtoreq. | Superscripted greater to or equal |
| +152 | | 1/8 | |
| +153 | | .sub..gtoreq | Subscripted greater to or equal |
| +154 | | 3/8 | |
| +155 | ≈ | .lmtorsim. | less than or approx. equal to |
| +156 | | 3/8 | |
| +157 | ≈ | .gtorsim. | greater than or approx. equal to |
| +158 | | 5/8 | |
| +159 | | .sup..lmtorsim. | |
| +160 | | 5/8 | |
| +161 | | .sup..gtorsim. | |
| +162 | | 7/8 | |
| +163 | | .sub..lmtorsim. | |
| +164 | | 7/8 | |
| +165 | | .sub..gtorsim. | |
| +166 | | 1/3 | |
| +167 | | .perspectiveto. | Perspective to symbol |
| +168 | | 1/3 | |
| +170 | | 2/3 | |
| +171 | | .A. | Open bold bracket (reissue) |
| +172 | | 2/3 | |
| +173 | | .1. | Close bold bracket (reissue) |
| +174 | | 1/4 | |
| +176 | | 1/4 | |
| +L78 | | 1/2 | |
| +180 | | 1/2 | |
| +182 | | 3/4 | |
| +183 | | 3/4 | |
| +184 | | 3/4 | |
| +185 | ≡ | .quadbond. | Quadbond symbol (atomic/chemical) |
| +186 | | .sup..crclbar | Superscripted neg. earth symbol |
| +187 | | .sub..crclbar. | Subscripted neg. earth symbol |
| +188 | | .sup..sym. | Superscripted pos. earth symbol |
| +189 | | .sub..sym. | Subscripted pos. earth symbol |
| +190 | # | | 23 Number Sign |

| | | | |
|------|--------------|------------------|-----------------------------------|
| +191 | ✧ | .notlessthan. | not less than symbol |
| +192 | ✧ | .notgreaterthan. | not greater than symbol |
| +193 | ♩ | .music-natural. | Music natural sign |
| +194 | → | .apprch. | Approaches symbol |
| +195 | ⊥ | .perp. | Perpendicular symbol (uptack) |
| +196 | $\sqrt[4]{}$ | .fourthroot. | Fourth root symbol |
| +198 | ™ | .TM. | Trademark symbol |
| +199 | SM | .SM. | Service mark symbol |
| +200 | ® | .RTM. | Registered trademark symbol |
| +201 | © | .COPYRG. | Copyright symbol |
| +202 | | .sup.s | Superscripted apostrophe |
| +203 | | .sub.' | Subscripted apostrophe |
| +204 | | .sup." | Superscripted double prime |
| +205 | | .sub." | Subscripted double prime |
| +206 | | .sup.' | Superscripted prime |
| +207 | | .sub.' | Subscripted prime |
| +208 | | .sup..degree. | Superscripted degree symbol |
| +209 | | .sub..degree. | Subscripted degree symbol |
| +210 | | .sup.— | Superscripted en dash |
| +211 | | .sup.— | Superscripted em dash |
| +212 | | .sub.— | Subscripted en dash |
| +213 | | .sub.— | Subscripted em dash |
| +214 | | .sup..div. | Superscripted division sign |
| +215 | | .sub..div. | Subscripted division sign |
| +216 | | .sup..infin. | Superscripted infinity symbol |
| +217 | | .sub..infin. | Subscripted infinity symbol |
| +218 | | .sup.' | Superscripted single open quote |
| +219 | | .sup.' | Superscripted single close quote |
| +220 | | .sup..dblquote. | Superscripted double open quote |
| +221 | | .sup..dblquote. | Subscripted double close quote |
| +222 | | .sub.' | Subscripted open single quote |
| +223 | | .sub.' | Subscripted close single quote |
| +224 | | .sub.dblquote. | Subscripted double open quote |
| +225 | | .sub.dblquote. | Subscripted double close quote |
| +226 | | .sup.1 | Superscripted script 1 |
| +227 | | .sub.1 | Subscripted script 1 |
| +228 | | .sub.\$ | |
| +229 | | .sup.# | |
| +230 | | .sub.< | NOTE: Column 1 numbers 228 |
| +231 | | .sup.< | thru 253 represent |
| +232 | | .sub.> | special characters as |
| +233 | | .sup.> | previous defined but |
| +234 | | .sub..vertline. | appearing either subscripted or |
| +235 | | .sup..vertline. | superscripted. |
| +236 | | .sup..not <. | |
| +237 | | .sub..not <. | |
| +238 | | .sup..not >. | |
| +239 | | .sub..not >. | |
| +240 | | .sup..noteq. | |
| +241 | | .sub..noteq. | |
| +242 | N/A | .sup. | (never assigned) |
| +243 | N/A | .sub. | (never assigned) |
| +244 | N/A | .sup. | (never assigned) |

+245 N/A .sub. (never assigned)
 +246 .sup..sqrt.
 +247 .sub..sqrt.
 +248 .sup..cuberoot.
 +249 .sub..cuberoot.
 +250 .sup..function.
 +251 .sub..function.
 +252 .sup..4th root.
 +253 .sub..4th root.

NOTE: Column 1 numbers 261 thru 286 represent the alphabet in "**small caps**". They will appear in APS as normal caps (as well as in MAPS format)

+261 A
 +262 B
 +263 C
 +264 D
 +265 E
 +266 F
 +267 G
 +268 H
 +269 I
 +270 J
 +271 K
 +272 L
 +273 M
 +274 N
 +275 O
 +276 p
 +277 Q
 +278 R
 +279 S
 +280 P
 +281 U
 +282 V
 +283 W
 +284 X
 +285 Y
 +286 Z

+290 .sbsb.0
 +291 .sbsb.1
 +292 .sbsb.2
 +293 .sbsb.3
 +294 .sbsb.4
 +295 .sbsb.5
 +296 .sbsb.6
 +297 .sbsb.7
 +298 .sbsb.8
 +299 .sbsb.9

NOTE: Column 1 numbers 290 thru 299 represent numeric values 1 - 9 as being a subscript of a subscript.

+300 .sup..alpha.
 +301 .sub..alpha.
 +302 .sup..omega.
 +303 .sub..omega.
 +304 .sup..theta.
 +305 .sub..theta.
 +306 .sup..lambda.
 +307 .sub..lambda.

NOTE: Column 1 numbers 300 thru 337 represent special characters as previously defined but appearing either subscripted or superscripted.

+308 .sup..eta.
 +309 .sub..eta.
 +310 .sup..mu.
 +311 .sub..mu.
 +312 .sup..delta.
 +313 .sub..delta.
 +314 .sup..pi.
 +315 .sub..pi.
 +316 .sup..gamma.
 +317 .sub..gamma.
 +318 .sup..SIGMA.
 +319 .sub..SIGMA.
 +320 .sup..epsilon.
 +321 .sub..epsilon.
 +322 .sup..phi.
 +323 .sub..phi.
 +324 .sup..sigma.
 +325 .sub..sigma.
 +326 .sup..nu.
 +327 .sub..nu.
 +328 .sup..tau.
 +329 .sub..tau.
 +330 .sup..beta.
 +331 .sub..beta.
 +332 .sup..upsilon.
 +333 .sub..upsilon.
 +334 .sup..psi.
 +335 .sub..psi.
 +336 .sup..rho.
 +337 .sub..rho.
 +338 .sbsb.a
 +339 .sbsb.b
 +340 .sbsb.c
 +341 .sbsb.d
 +342 .sbsb.e
 +343 .sbsb.f
 +344 .sbsb.g
 +345 .sbsb.h
 +346 .sbsb.i
 +347 .sbsb.j
 +348 .sbsb.k
 +349 .sbsb.l
 +350 .sbsb.m
 +351 .sbsb.n
 +352 .sbsb.o
 +353 .sbsb.p
 +354 .sbsb.q
 +355 .sbsb.r
 +356 .sbsb.s
 +357 .sbsb.t
 +358 .sbsb.u
 +359 .sbsb.v
 +360 .sbsb.w

NOTE: Column I numbers 338 thru
 363 represent lower case
 letters that appear as a
 subscript of a subscript.

+361 .sbsb.x
 +362 .sbsb.y
 +363 .sbsb.z
 +364 .sbsp.a
 +365 .sbsp.b
 +366 .sbsp.c
 +367 .sbsp.d
 +368 .sbsp.e
 +369 .sbsp.f
 +370 .sbsp.g
 +371 .sbsp.h
 +372 .sbsp.i
 +373 .sbsp.j
 +374 .sbsp.k
 +375 .sbsp.l
 +376 .sbsp.m
 +377 .sbsp.n
 +378 .sbsp.o
 +379 .sbsp.p
 +380 .sbsp.q
 +381 .sbsp.r
 +382 .sbsp.s
 +383 .sbsp.t
 +384 .sbsp.u
 +385 .sbsp.v
 +386 .sbsp.w
 +387 .sbsp.x
 +388 .sbsp.y
 +389 .sbsp.z
 +390 .spsp.0
 +391 .spsp.1
 +392 .spsp.2
 +393 .spsp.3
 +394 .spsp.4
 +395 .spsp.5
 +396 .spsp.6
 +397 .spsp.7
 +398 .spsp.8
 +399 .spsp.9
 +400 .sbsp.0
 +401 .sbsp.1
 +402 .sbsp.2
 +403 .sbsp.3
 +404 .sbsp.4
 +405 .sbsp.5
 +406 .sbsp.6
 +407 .sbsp.7
 +408 .sbsp.8
 +409 .sbsp.9
 +410 .spsb.0
 +411 .spsb.1
 +412 .spsb.2
 +413 .spsb.3

NOTE: Column 1 numbers 364 thru 389 represent lower case letters that appear as a subscript to a superscript.

NOTE: Column 1 numbers 390 thru 399 represent numeric value 1 - 9 as being a superscript of a superscript.

NOTE: Column 1 numbers 400 thru 409 represent numeric values 1 - 9 as being a subscript of a superscript.

NOTE: Column 1 numbers 410 thru 419 represent numeric values 1 - 9 as being a superscript

+414 .spsb.4 of a subscript.

+415 .spsb.5

+416 .spsb.6

+417 .spsb.7

+418 .spsb.8

+419 .spsb.9

+438 .spsp.a

+439 .spsp.b

+440 .spsp.c

+441 .spsp.d

+442 .spsp.e

+443 .spsp.f

+444 .spsp.g

+445 .spsp.h

+446 .spsp.i

+447 .spsp.j

+448 .spsp.k

+449 .spsp.l

+450 .spsp.m

+451 .spsp.n

+452 .spsp.o

+453 .spsp.p

+454 .spsp.q

+455 .spsp.r

+456 .spsp.s

+457 .spsp.t

+458 .spsp.u

+459 .spsp.v

+460 .spsp.w

+461 .spsp.x

+462 .spsp.y

+463 .spsp.z

+464 .spsb.a

+465 .spsb.b

+466 .spsb.c

+467 .spsb.d

+468 .spsb.e

+469 .spsb.f

+470 .spsb.g

+471 .spsb.h

+472 .spsb.i

+473 .spsb.j

+474 .spsb.k

+475 .spsb.l

+476 .spsb.m

+477 .spsb.n

+478 .spsb.o

+479 .spsb.p

+480 .spsb.q

+481 .spsb.r

+482 .spsb.s

+483 .spsb.t

+484 .spsb.u

NOTE: Column 1 numbers 438 thru 463 represent lower case letters that appear as superscript of a superscript

NOTE: Column 1 numbers 464 thru 489 represent lower case letters that appear as superscript of a subscript.

| | | |
|------|---------|--------------------------------------------------|
| +485 | .spsb.v | |
| +486 | .spsb.w | |
| +487 | .spsb.x | |
| +488 | .spsb.y | |
| +489 | .spsb.z | |
| +500 | • | .cndot. Center Dot (Solid Dot) |
| +501 | ● | .circle. Large Circle (filled black) |
| +502 | ⊙ | .circleincircle. Circle in a large circle |
| +503 | ◐ | .lhalfcircle. Left-half circle -hollow & filled |
| +504 | ◑ | .rhalfcircle. Right-half circle -hollow & filled |
| +505 | ⊘ | .dottedcircle. Dotted circle |
| +506 | ⎵ | .THorizBrace. Top horizontal brace |
| +507 | ⎴ | .BHorizBrace. Bottom horizontal brace |
| +508 | ◐ | .dotlhalfcircle. Dotted left-half circle |
| +509 | ◑ | .dotrhalfcircle. Dotted right-half circle |
| +510 | ◐ | .dotthalfcircle. Dotted top-half circle |
| +511 | ◑ | .dotbhalfcircle. Dotted bottom-half circle |
| +512 | ◐ | .solthalfcircle. Solid top-half circle |
| +513 | ◑ | .solbhalfcircle. Solid bottom-half circle |
| +515 | ♂ | .male. Male symbol |
| +516 | ♀ | .female. Female symbol |
| +517 | § | .sctn. Section symbol |
| +518 | ∅ | .0. Slashed zero (light print) |
| +519 | ∅ | .0. Slashed zero (bold print) |
| +520 | ⊎ | .orgate. or gate symbol (logical OR: V) |
| +521 | Å | .ANG. Angstrom (upper case) |
| +522 | å | .ang. Angstrom (lower case) |
| +601 | ◦ | .smallcircle. Small circle |
| +701 | ◯ | .largecircle. Large circle |

The following special characters precede data on the patent full-text/APS file:

| <u>Column 2</u> | <u>Description</u> |
|-----------------|------------------------------|
| .sub. | Subscript |
| .sup. | Superscript |
| .sbsb. | Subscript of a Subscript |
| .spsp. | superscript of a superscript |
| .sbsp. | subscript of a superscript |
| .spsb. | superscript of a subscript |

Subscripted and supercripted data will be terminated by a space, a comma (,), a period (.), a right parenthesis ()) or another subscript .sub. or superscript .sup..

NOTE 1 - The following defines the appearance of original and added data in a Reissue Patent on the Patent Full-Text APS File.

Original data, that forms no part of the reissue patent, will begin with .~. and end with .]. New data (appears Italicized on the printed document) will begin with .Iadd. and end with .laddend.

NOTE 2 - The following defines the appearance of original and added data in a Reissue of a Reissue Patent on the Patent APS file.

Original data of the first reissue patent, that forms no part of the current reissue patent, will begin with .~. .(. and end with .]. .J. New data of the first reissue patent (appears in BOLD print on the current printed document) will begin with .Badd. and end with .Baddend.

NOTE 1 above applies to the appearance of original and added data on the current Reissue Patent.

We started a project to convert all of the Dot Codes to HTML and UTF-8 characters, but, as we proceeded, we discovered that it would not be a simple substitution exercise. Below is one example of hundreds we ran across.

Examples of why automated conversion of all Dot Codes is more difficult than we first thought.

Example of valid Code: **.En.**

From US4231672A

NAM Blanpain; Guy

CTY Verneuil.**En.**Halatte

CNT FRX

.En. found but is NOT a Dot Code in US4238984A:

38 at the timing T.sub.**En.** In this case, the output Q.sub.1, Q.sub.2 and

The dots are NOT shared – we handled this type of exception with a simple change to our code.

.En. found again, but is NOT a Dot Code in US3992364A:

F E.As. No Illness 250 245

M O.**En.** Cerebral Sclerosis

The O.En. is an abbreviation used in biology, and this is not nearly as easy in programming logic.

We will continue with the project until they are completed.

APPENDIX EP – EPO Specific Internal Tags

This is the listing of Internal **B000EP** tags appearing in MAPS EP documents as comments (called **<eptags>** in XML data). This data does not generally appear on printed EP publications, and it is not supported as indexed data in any of our systems, but we do use one of the tags to help identify EuroPCT files (**B003EP**).

We keep the data in the MAPS files since we prefer to error on the side of inclusion where data is concerned for our customer's benefit (you may know things we don't), and you may have a use for them – so we keep them for you.

For completeness, we keep all of these in MAPS Comment Fields exactly as the as the EPO published them in the data (extra spaces and all). For example, this one:

CMNT B008EP: J

| EP000 Tag | Reported Use | Uses these tags: | Used by this tag: |
|-----------|---------------------------------------------------------|---------------------------------------------------------|-------------------|
| B001EP | Select. mask for states involved | | eptags |
| B002EP | Changes to bibliographic data | ep-chg-info | eptags |
| B003EP | Indicator that application was made with PCT to the EPO | Indicates that a Kind A Pub may not be published by EPO | eptags |
| B004EP | Re-establishments of rights indicator | | eptags |
| B005EP | Printer id. | | eptags |
| B006EP | Ind. for inter. applic. | | eptags |
| B007EP | Reserved | | eptags |
| B008EP | Small change indicator | | eptags |
| B009EP | Transl. Of B725EP for the EP Bulletin | | eptags |
| B010EP | Other rights and legal means of execution | B011EP | eptags |
| B011EP | Serial number date and states | date, dnum, ctry | B010EP |
| B015EP | Number of documents to be printed | | eptags |
| B020EP | Biblio rec. creation | | eptags |
| B021EP | Biblio. Rec. correction | | eptags |
| B030EP | Legal status data | | eptags |
| B050EP | Free text data | B051EP, B052EP | eptags |
| B051EP | Language | | B050EP |
| B052EP | Free text | | B050EP |
| B053EP | Remarks | | eptags |
| B060EP | Check data | | eptags |
| B061EP | Data relating to fees | | eptags |
| B065EP | Data relating to fees | | |
| B070EP | B pub. technical field | | eptags |
| B078EP | Date of 'no oppo. filed' | Date | eptags |
| B080EP | | B081EP, B082EP, B083EP | eptags |
| B081EP | Inspection place | | B080EP |

| | | | |
|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|--------|
| B082EP | EPO code relation | tll, fmm, snm, sfx, iid, irf, onm, syn, oid, odv, did, adr, dtxt, retry, nctry | B080EP |
| B083EP | Free text | | B080EP |
| B090EP | B0900EP contains each instance of a limitation with B091EP and B093EP containing the rest of the fields as dates and limitation description details (numerically coded). This information is from an EPO internal database and is not contained on the printed publication. | B0900EP, B091EP, B0911EP, B0912EP, B0913EP, B0914EP, B093EP, B0931EP, B0932EP, B0933EP | Eptags |
| B130EP | We're still researching this one... | | |
| | | | |

If someday, corresponding PCT B000WO fields exist in PCT publications, they will be handled the same way (and you will get the data).

Appendix H

USPTO, EPO and WIPO Patent-Data Related History

This appendix provides a brief run-down on patent-related data and its acquisition, from our point of view, which is the view of someone who directly interfaces with the various patent authorities to find, purchase, reformat, and make data available to others. Please keep in mind that how we see certain events or policies and how they are viewed from within a patent authority will surely vary.

THE USPTO, APS and the 1970s and 80s

USPTO Data – the Early days. **APS** was an acronym for the **USPTO's Automated Patent System**. It is our opinion (shared by many others) that it was one of the best designed systems of its day, and the only one we are aware of functioning as early as 1971. Without comparing it to early banking systems, I can't think of another system that processed combined graphics and text oriented data that lasted and was used for almost 30 continuous years performing its mission (and it did it well – very well). The patent data storage format was documented in what was known as the “Green Book” (**GB**).

The system ran on IBM or IBM compatible mainframe computers (e.g., Amdahl systems). The USPTO search systems used it, the system dumped data onto 200MB IBM style cartridge tapes that were sold to the public (it was expensive, too – about \$250,000 for a complete set of the data).

The older image data was stored in “Yellow Book” format which was basically CCITT G4 compressed image data stored in 20 kilobyte chunks that could be handled by the mainframe. It was quite impressive at the time. Their search and display systems were able to pull this data from storage devices for display in the PTO for examiners, and eventually end users. We created custom software to convert the 20KB chunks of image data into the single page TIF files, then on to Multi-page TIFFs (Version 6 of the TIF spec date 1993) which are in use today in most systems storing patent data, including the USPTO's new systems.

The APS system also handled the patent text which was stored in groups of 80 byte fixed length records using the IBM EBCDIC character encoding system. The 80 byte record was a descendant of the “Punch Card” – 80 holes across, and 8 or more down – or places for holes, anyway. The original “Green Book” (GB) specification documented the byte records understood by APS, along with an ingenious method of storing math and scientific symbols along with letters in the Greek alphabet and other character sets, and it also allowed for various character styles such as superscript, subscript, underlined, etc. The system was known as **Dot Codes**.

Dot Codes are word based where a character such as the Greek letter **Beta** would be replaced by the word BETA (upper case for Capital Beta, lower for small beta) surrounded by dots (periods) like this: **.beta**.

Styles such as subscript were prefixed another Dot Code like this: **.sub**. Technical terms in genetics or chemical strings like **endo- α -1** or **6- β -glucanase** were stored as **endo-.alpha.-1** and **6-.beta.-glucanase**. The entire string could be indexed in search systems or individual words without the punctuation which meant you could search it with simple phrases like this **endo alpha 1** and **6 beta glucanase** which would turn up a patent containing this reference:

*Biely, P. et al., "Purification and properties of alternanase, a novel **endo-.alpha.-1**, 3-60 -1, 6-. beta.-*

glucanase", European Journal of Biochemistry, (1994), vol. 226, No. 2, pp. 633-639. cited by other.

This text based "Dot Code" encoding system is still in use today (2015) at the USPTO in their search systems and in older (pre-2000) US MAPS Description and Claims text. IP Data Corporation took on the task of converting all Dot Codes to UTF-8 character codes, but there have been several delays due to certain ambiguities we have run across which requires us to pull up the images copies for over 100,000 documents to ensure we have the correct codes. It is better to leave Dot Codes than make assumptions that leaves us with less than 100% accurate data. If you need a copy of Appendix C of the Green Book, ask us.

The USPTO replaced the GB data in 2000 with the XML data, and they **convert backwards** to the GB dot codes data for their online systems. We did the same thing to keep them compatible with the 25 years of Pre-2000 data. But we have already started a project to convert all of the GB raw data and Dot Codes into UTF-8 characters and HTML, and it will continue until completed.

If you used data in the early 2000s, we're sure you noticed the HTML entities "creeping" into the MAPS files as many came along with no translation to the Green Book Appendix C system, and you may also have noticed many showing up as properly displayed UTF-8 characters in your browsers.

The late 70s and the EPO and WIPO Get Involved

In 1978 the very first World Patent applications were published by The World Intellectual Property Organization (**WIPO**) as directed under the Patent Cooperation Treaty (**PCT**) that was signed by many countries in 1970. It provided a way to use a standard application and submit it to more than one patent authority.

The EPO also published their first applications in 1978, and unlike the WIPO, they had Grant and enforcement authority for EU member states (countries). **In 1980, they granted the first EP patents.**

As part of WIPO's mission, it was to produce patent data storage, formatting and publishing standards through a cooperative process with all of the treaty signatories. And **WIPO** did produce several good standards defining Document Kinds, minimum document bibliographic sections, and how to mark up and identify the "front page" data fields so they were easy to identify. The EPO was the first to adopt many of these standards such as assigning the proper recommended Kind Codes to various documents (A1, A2, A8, A9, B1, B2, etc.). WIPO, however, was not so fast to adopt many of its own standards. But WIPO finally caved in and began using the proper correction document Kind Codes in 2009.

The first data available for general public consumption for WIPO and EPO data was supplied and managed by the EPO for both authorities, and consisted of OCR text from their EPOQUE Text database and CCITT G4 compressed collections of single page images for each document distributed via CD ROM by subscription, and also as CD-ROM Back-file collections. Larger distributions of other formats were available on tape from BACON (a **BACK**-file **CON**version system using facsimile images) which was a complex, multi-part system of capture, conversion, formatting and storage of WIPO (and other) data.

In the mid 1990s as SGML type tag based text storage formats were becoming popular, the USPTO jumped on the bandwagon to find a good system to replace their APS system (reaching a state of decrepitude, at that point – but still admired by most everyone that used it). In 1996 they decided upon an early SGML variant and after several major changes over a 4 to 5 year period, we ended up with XML in a very similar format to what we now get from the USPTO.

In the early 2000s, the USPTO undertook a major project to OCR all of the text of US patents prior to 1979 all the way back to 1840. Why 1979 I don't know – we have great text for 76 to 79. It is available at no cost from various USPTO distribution systems. They have performed absolutely no cleanup work on them – it is raw OCR data from mediocre to poor paper copies. ***Caveat Emptor!***

MAPS was originally designed for US data, and it was loosely based on the USPTO APS “Green Book” specification with the computer “mainframe eccentricities” removed and cleaned up. In other words, it went from an 80 column fixed length, line-oriented format to a text-file based, paragraph oriented format where the logical line was a paragraph and the same as the physical line (e.g., reaching the physical EOL also means the end of the logical data for that field or text paragraph).

Several large corporations as well as several professional patent search firms use the MAPS format that we supply, some with our P9 Patent Search system, while others have built their own systems to use within their organizations. Two years ago we were asked to use MAPS as a basis for a new XML product since a lot of cleanup work had been done on older PCT publications and since it was a single format for the supported authorities, the XML would also be a single format greatly reducing the work needed for users to create various types of multiple authority database.

Appendix K – Kind Codes, All Authorities

Kind Codes for PCT and EP Documents

WIPO and the **EPO** have been “loosely” following the ST.16 recommendation for Kind Codes for quite some time and consequently, the Kind Codes have become an integral and important part of the **mostly** unique document numbers that describe the WIPO and EP documents. The USPTO changed kind codes in 2001 to more closely follow WIPO ST.16 recommended standards. Below are the Document Kind Codes used in WO and EP publications.

Kind Codes for PCT and EP Application Related Publications

| | | |
|----------|-------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| A1 | Patent Application w/Search Report | |
| A2 | Patent Application Without Initial Search Report | An A3 Document with the Front page information and the pages of the Search Report* can follow almost immediately, or can be several years behind. * International Search Report in the case of WIPO Applications. |
| A3 | Search Report for A2 Document | Previously used codes B1 and B2 are now used for granted Patents (EPB only) |
| A4 A4 | EP Supplemental Search Reports WO Amended Claims | EP does not publish Amended Claims, and <i>Supplemental Search Reports</i> are not available in Bulk Data |
| B1 | Amended Claims (WO prior to 2009) | Why did they use B1? We have NO idea... |
| A8 | Front page (Bib) corrections | This usually causes the front page to be republished. |
| A9 | Major Correction (entire doc) | The Entire Application Document is replaced. |

Kind Codes for Granted Patent and Patent Related documents for **EP-“B” documents, only** (WIPO only publishes applications under the PCT).

| | | |
|----|--------------------------------------------|-------------------------------------------------------------------------------------------------------|
| B1 | Granted Patent | Without published applications |
| B2 | Granted Patent | With previously published applications |
| B3 | Granted Patent (with Limitations) | These are granted patents with limitation decisions. Generally with previously published applications |
| B8 | Front page corrections on Grant (Bib data) | This causes the front page to be republished. |
| B9 | Correction to Grant text or Figures. | The Entire Document is replaced. |

Kind Codes for US Patent Documents

Kind Codes Assigned to Published Documents (and included in data files)

NOTES:

- “n” (in table below) represents a value 1 through 9 denoting the publication sequence or level.
- Many of these publications are only available in the Application file histories or upon request from the USPTO.

| Kind Code | Publication |
|------------------|--------------------|
|------------------|--------------------|

| | |
|-----------|---------------------------------------------------------------------------------------------|
| A | 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 | Correction published Utility Patent Application |
| Bn | Reexamination Certificate issued prior to Jan 2, 2001. |
| B1 | Utility Patent Grant (no pre-grant publication) issued on or after January 2, 2001. |
| B2 | Utility Patent Grant (with pre-grant publication) issued on or after January 2, 2001. |
| Cn | Reexamination Certificate issued on or after January 2, 2001. |
| Fn | Supplemental Examination Certificate published after Sept.16, 2012. |
| Jn | Post Grant Review Certificate published after September 16, 2012. |
| Kn | Inter Partes Review Certificate published after September 16, 2012. |
| On | Derivation Certificate published after March 16, 2013. |
| E1 | Reissue Patent issued January 2, 2001 and Later |
| E | Reissue Patent issued prior to January 2, 2001 |
| Bn | Reexamination Certificate of a Reissue Patent. Certificate issued prior to January 2, 2001. |
| Cn | Reexamination Certificate of a Reissue Patent. Certificate issued on or after Jan. 2, 2001. |
| | SIRs began with the Dec 3, 1985 issue. |
| H1 | Statutory Invention Registration (SIR) Docs. Published on or after January 2, 2001. |
| H | Statutory Invention Registration (SIR) Docs. SIRs Prior to January 2, 2001 |
| I1 | "X" Patents issued from July 31, 1790 to July 13, 1836. |
| I2 | "X" Reissue Patents issued from July 31, 1790 to July 13, 1836. |
| 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 (very old) |
| Not used | Non-Patent Literature |
| P1 | Plant Patent Grant issued on or after January 2, 2001 |
| P | 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 pre-grant publication) issued on or after January 2, 2001. |
| P3 | Plant Patent Grant (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 2001 and later |
| S | Design Patent Grant Issued Prior to January 2, 2001 |

Appendix L – EP and WO Languages in MAPS

EP and WO publications are published in several languages as approved by the European Patent Convention (EPC) for **EPs**, and the Patent Cooperation Treaty (PCT) for **WOs**.

The MAPS specification defines different codes for the four basic text sections which **allows us to define several different languages in the same document, simultaneously**, without adding some other mechanism (such as XML attributes) to define the language in that section.

The following table lists the MAPS codes for the sections and languages:

| Section VVVV | Language >> | English (en) or *PBLN defined | German (de) | French (fr) | Other (WOs only) PBLN defined |
|--------------------|-------------|----------------------------------|-------------|-------------|----------------------------------|
| Title | | TITL | TITD | TITF | TITO |
| Abstract | | ABST | ABSD | ABSF | ABSO |
| Description | | *DESC | DESD | DESF | N/A |
| Claims | | *CLMS | CLMD | CLMF | N/A |

The MAPS codes are used to define languages in the text sections following **THREE** basic rules:

1. In **EP Publications**, the Codes ALWAYS define the Language listed in the Table (i.e., **DESC** and **CLMS** are always English, and all other codes are as listed in the table, as well (*very easy*)).
2. In **ALL Publications OTHER than EPs**, the **PBLN** listed language defines the language in the **DESC** and **CLMS** sections (this covers U.S. Publications, and also covers up to 30 languages in PCT documents).
3. When **TITO** or **ABSO** is used in **WOs**, the language is the language specified in the **PBLN** Code (Publication Language). This allows us to easily list the Titles and Abstracts in PCTs in English, French, and the OTHER (native) language listed by the **PBLN** Code (the Published Language).

EP Application Languages

In EP Applications, the Title, Abstract, Claims and Description will be in one of the three official EPC languages: **English, German or French**. The Abstract, Description and Claims text sections will be in the single language using the codes from a single column in the above table for that language, plus, EP Applications always include the Title in all three approved languages using the codes from the Title Row in the table.

EP Granted Patent Languages

EP-B publications (granted patents) contain the Title in all three languages with codes **TITL, TITD** and **TITF**, the Claims Text in all three languages with codes **CLMS, CLMF** and **CLMD**, while the Descriptive Text will be in the Language of the Application and will use one of the three codes **DESC, DESD** or **DESF** to define

it.

The **DESD (German)** and **DESF (French)** Description Codes, and the **CLMD (German)** and **CLMF (French)** Claims Text Codes are only used in **EP-A** and **EP-B** publications and this specification as of Version 6.3 guarantees that the **DESC** (Description) and **CLMS** (Claims) sections in **Non-EP** publications will contain the language listed in the **PBLN** field. *In other words*, we STOPPED using the **DESD**, **DESF**, **CLMD** and **CLMF** codes in all publications OTHER than EPs, and this will not change.

WO (PCT) Application Languages

With PCT applications there are literally dozens of languages allowed for the Description and Claims text sections, but the Title and Abstract are still listed in English (**ABST**) and in French (**ABSF**) plus the Native language using **TITO** and **ABSO** which are defined by the language in the **PBLN** code.

When the publication language is something other than English, like Japanese (**ja**), Korean (**ko**), a Chinese dialect (**zh**) or Russian (**ru**), as examples, the Bibliographic data elements will still be in English, plus, the Title and Abstract are also usually supplied in French, as well, plus, the **TITO** and **ABSO** fields will always be used if these data are available, which is not always the case. Sometimes they are simply not supplied.

The current ST.9 Standard indicates the language in the field is a two letter language from **International Standard ISO 639:1988**. ISO tells us that the NEW standard is **ISO 639-1:2002**. **And they want about \$140 for a copy****, so I recommend you use the included list in this document in **APPENDIX B** titled **Country and Language Codes**.

You will find the correct language in the **PBLN** field (***NOT the AFLN field***) for the publication. Also, don't look at the country of origin from the application numbers (e.g., the country is part of the number) and assume, for example, that a JP Application number means the text is in Japanese (ja). Quite often it isn't. We are now receiving many **Japanese PCTs in English** that are NOT OCR sourced.

** We broke down and bought a copy, and I'm sorry we did since it was a waste of funds because the partial copy we found online was accurate, and although not complete (it's a subset) it has ALL of the PCT languages supported by WIPO.

Appendix M – Math-ML Data

Math-ML is a tag-based method for storing and describing Math formulas that can be accurately rendered using a browser. However, most browsers cannot display Math-ML formulas without some modifications or add-ons.

Typical Browsers and Math-ML

Not all browsers are created equal, but our testing shows that some are more equal than others (or so the old saying goes). Our tests with Math-ML data from the patent data as our only source, since that is our primary interest, have provided the following results as far as which browsers can display Math-ML as delivered (or downloaded with default capabilities).

Browsers That *Can't* Display Math-ML Without Modifications

The following browsers **CANNOT display Math-ML** data as delivered in the patent data:

- Internet Explorer 9, 10, or 11.
- Chrome Versions 44.x, 45.x and 46.x

Browsers That *CAN* Display Math-ML Without Modifications

The following browsers **CAN display Math-ML** data as delivered except that several often used characters are displayed incorrectly making the formulas very difficult to read (the reason is discussed below):

- Firefox Versions 39.x, 40.x and 41.x
- SeaMonkey Versions 2.2x to 2.38 (same rendering engine as Firefox)
- Opera 32.0 (still being updated – a actually a nice browser)
- Safari Version 5.1.7

Math-ML Files included as Separate HTML-5 files

If you require the display of formulas contained in the patent data, we have created individual HTML-5 containers and saved the Math-ML data from the patents and applications in these containers and named them using the Math-ML call-out numbers in the patents, The references to the Math formula number remain in the patents for you to use as you build your HTML page.

The Math-ML HTML files look like this:

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="utf-8">
  <title>US8930150B2 MATH-US-00002</title>
</head>
<body>
  <maths id="MATH-US-00002" num="00002">
```

```
>>> FORMULA DATA IN THE MATH-ML LANGUAGE <<<
</math>
</body></html>
```

The MathML files are stored in a directory named **mathml** inside the weekly patent and application issues. If you store these HTML files with the APS or XML files as delivered, and you store them in the year/date path (or year/date path for Windows) like the MAPS data is stored, creating the path to them will be almost the same as opening one of the MAPS or MAPS-XML files.

Using a Generic Inline frame to display Formula Files

The HTML `<iframe>` command is very easy to use to display the formulas because it can be set to automatically set the height to match the height of the formula. This is accomplished by using the CSS style inline with the command as follows:

```
<iframe src="MapsURLroot/2015/20150106/mathml/US8929637B1-MATH-US-00005.html"
        style="border:none; width:100%"></iframe><br>
```

Using CSS and setting the style to **border:none** and **width:100%** automatically sets the height of the frame to match the height of the displayed formula, and sets the width for the page width, or the width of the container it is in, such as a table cell or `<div>` `</div>` with the width you set up in the CSS file.

Why We Pulled the MathML From the Files

If you display the MathML with any of the Browsers listed above as NOT supporting the mathML, the display can look pretty nasty and will even overwrite some of the text above or below the improperly displayed formula data. Browsers are supposed to “gracefully” ignore what they can't display correctly, but that is not always the case.

By pulling the MathML out and placing the data in separate HTML files, **you have the option** of calling up the files for display if desired, or just ignoring the MathML altogether and telling users to see the Image files. We tried to leave in whole words and variables for indexing and as much of the formulas as possible without ruining your display of the remaining data.

Building the Math-ML Filename

The Call-Out format **is not a true file reference**, but instead is the usual MAPS notification for a formula telling you to see the facsimile image. The difference is that it is numbered (with the USPTO numbering scheme), and we use that number to build the filename. Here is an example of the first Formula in US Application **US20150000145A1** published January 1st, 2015:

[Math-ML: **MATH-US-00001** See Facsimile image]

The MathML HTML file is named: **US20150000145A1-MATH-US-00001.html**

You start with the **Publication Number**, add a **dash** (2D hex), then **add the MathML number** (highlighted in yellow above) and finally **append “.html”** onto it. This is the complete filename for the MathML HTML

file. If your mathml directory is stored with the MAPS-XML files, or MAPS APS files, in each weekly directory, use the SAME path you used to open the MAPS patent file, but add mathml to it BEFORE you append the filename you built, above. You will end up with something like :

“<http://192.168.1.37/mapsxml/2015/20150106/mathml/US20150000145A1-MATH-US-00001.html>”

MathML with Default Math Fonts on Firefox, Safari and Opera

As described above, we have started saving the MathML formulas in HTML5 compatible files with very minimal headers, as one file per formula, or formula group. The MathML files contain the exact MathML code as provided by the USPTO, and they are saved **before** we convert the Ampersand Entities to UTF-8 Binary character codes. Our first round of tests were completed with the UTF-8 binary versions but it was more difficult for us to lookup the Unicode characters in the individual HTML files we built, so we changed our process to save the files BEFORE we completed the character code conversions to binary UTF-8. The browser doesn't care which way we use them, and when we get time we want to find substitutions for more characters to make the HTML files work even better without forcing anyone to modify their browsers to see the MathML formulas.

If you decide to display the MathML formulas in MAPS or MAPS-XML text for your users, they will require a browser with the MathML Rendering engine built-in or a browser with the proper extension/add-on required to display MathML. The browsers we have tested are listed above in the first section of this Appendix.

A Bug in Mozilla and Opera Browsers with the Default MathML Font

If your users run Firefox, Seamonkey or Opera, they may complain about characters that do not display correctly. They are no doubt using the default fonts the browsers were designed to use to render MathML characters. We have loaded several fonts that are recommended but none of the seem to fix the issues we are seeing in these browsers, so we investigated which characters are causing the problem. Here is what we found.

Safari Browsers (Mac and PC both) look OK Except for One Character

Your Mac users will be OK with the Default Font they use for MathML except there seems to be one character that does not display correctly (you get the typical empty box). We can't figure out why the default does not work. The correct character is in the font they use (Stix font that ships with OS X 10.7), but the rendering engine does not seem to select it correctly. But it's not one that seems to be used a lot so they will probably be OK and won't bug you about it (*too much...*).

Some MathML Uses Newer, Not-yet-supported Character Codes

We don't know why, but the designers of newer versions of MathML were not content to use certain previously defined math characters from the Unicode Math sets, even though they were the exact characters and symbols needed, **and previously used**. In these newer versions, they decided to assign their own codes to several very common Math symbols and Invisible operators using codes from the Unicode MathML private areas. **NIHS** maybe? (Not Invented Here Syndrome).

We got tired of seeing the strange boxes and Chinese characters, and we use the mathML all the time and you want a painless way to satisfy 95% of the problems for Windows/PC users, you can substitute seven of the most common “new” characters in the MathML data so the default Math font in Windows works as it once did for almost all of the MathML formulas found in the patent data. In Windows Mozilla Products and

with Opera, the default MathML font is the **Cambria Math** series.

We have a switch in our MAPS code converter that that does the substitution for the files we use in our local search systems, and we reversed ourselves again and **we have substituted the 7 most egregious character problems in the MathML data**. The character codes we substitute are identically defined and display correctly – they are simply earlier versions of the newer codes.

The UTF-8 binary values are 3 byte sequences, and are shown inside parenthesis in the below table that lists the seven characters:

| For Characters: | The USPTO sent: | The correct (older) substitute is: |
|---------------------|---------------------|------------------------------------------------------------|
| Invisible Times | - U+e89e (EE A2 9E) | but should be U+2062 (E2 81 A2) (this is used extensively) |
| Apply Function | - U+e8a0 (EE A2 A0) | but should be U+2061 (E2 81 A1) |
| Differential d | - U+f74c (EF 9D 8C) | but should be U+2146 (E2 85 86) |
| Exponential e | - U+f74d (EF 9D 8D) | but should be U+2147 (E2 85 87) |
| Imaginary i | - U+f74e (EF 9D 8E) | but should be U+2148 (E2 85 88) |
| Left Bracketing Bar | - U+f603 (EF 98 83) | but should be U+23a2 (E2 8E A2) (Left Bracket Extension) |
| Rite Bracketing Bar | - U+f604 (EF 98 84) | but should be U+23a5 (E2 8E A5) (Right Bracket Extension) |

The image directly below (or the top of the next page) are what users see in the Mozilla Browsers (Firefox, Seamonkey) **without the fix**, and the second image below is what the Opera browsers display **without the substitutions**.

The upper range Private Area characters will work (with a lot of work on your end), but when we use the more common Unicode values for the characters plus 2 of our own substitutes (the brackets) the browsers can use their default Math Fonts. **You no longer have to go through heroic efforts to make them work**. The Safari Browser only has one font issue, and it is usable as is.

Without them, here is what you see in the Mozilla Browsers with the **Cambria MathML Font**:

$$\begin{aligned}
 x_0 &= -\frac{a_4}{2 a_1} y_0 - \frac{a_5}{2 a_1} z_0 - \frac{a_7}{2 a_1} & (2 a) \\
 y_0 &= -\frac{2 a_1 a_6 - a_4 a_5}{4 a_1 a_2 - a_4^2} z_0 - \frac{2 a_1 a_8 - a_4 a_7}{4 a_1 a_2 - a_4^2} & (2 b) \\
 & (2 a_1 a_9 - a_7 a_5) (4 a_1 a_2 - a_4^2) - \\
 z_0 &= -\frac{(2 a_1 a_6 - a_4 a_5) (2 a_1 a_8 - a_4 a_7)}{(4 a_1 a_3 - a_5^2) (4 a_1 a_2 - a_4^2) - 2 (a_1 a_6 - a_4 a_5)^2} & (2 c)
 \end{aligned}$$

Next is what Opera looks like:

$$\begin{aligned}
 x_0 &= -\frac{a_4}{2 a_1} y_0 - \frac{a_5}{2 a_1} z_0 - \frac{a_7}{2 a_1} & (2 a) \\
 y_0 &= -\frac{2 a_1 a_6 - a_4 a_5}{4 a_1 a_2 - a_4^2} z_0 - \frac{2 a_1 a_8 - a_4 a_7}{4 a_1 a_2 - a_4^2} & (2 b) \\
 & (2 a_1 a_9 - a_7 a_5) (4 a_1 a_2 - a_4^2) - \\
 z_0 &= -\frac{(2 a_1 a_6 - a_4 a_5) (2 a_1 a_8 - a_4 a_7)}{(4 a_1 a_3 - a_5^2) (4 a_1 a_2 - a_4^2) - 2 (a_1 a_6 - a_4 a_5)^2} & (2 c)
 \end{aligned}$$

Yes, I think that's Chinese. **Next is what you see after the new characters were substituted using our StringSubMany utility** (We confirmed that it is correct by comparing it with the Facsimile image copy of the patents):

$$x_0 = -\frac{a_4}{2 a_1} y_0 - \frac{a_5}{2 a_1} z_0 - \frac{a_7}{2 a_1} \quad (2a)$$

$$y_0 = -\frac{2 a_1 a_6 - a_4 a_5}{4 a_1 a_2 - a_4^2} z_0 - \frac{2 a_1 a_8 - a_4 a_7}{4 a_1 a_2 - a_4^2} \quad (2b)$$

$$z_0 = -\frac{(2 a_1 a_9 - a_7 a_5)(4 a_1 a_2 - a_4^2) - (2 a_1 a_6 - a_4 a_5)(2 a_1 a_8 - a_4 a_7)}{(4 a_1 a_3 - a_5^2)(4 a_1 a_2 - a_4^2) - 2(a_1 a_6 - a_4 a_5)^2} \quad (2c)$$

With the MathML files we now ship, if your users are willing to use Firefox, SeaMonkey or Opera browsers, we think they will like the the MathML formula display quality.

Substituting the Older for the Newer Characters, If Needed

If you want to substitute your MathML files, download our **StringSubMany** package from the **SPECIAL** directory on the **FTP server** (discussed in the next section) and execute it like this:

```
C:\>StrSubMany X:\Path-to-MathML-Files\ *.html MathMLFontFix.txt <Enter>
```

StringSubMany Utility for Mass Text Substitution (Windows)

Our **StringSubMany** package is available in the **SPECIAL** directory on the **FTP server** and contains the following files:

| | |
|--------------------------|------------------------------------------------------------------------------------------|
| StrSubMany.exe | A Windows Executable Program |
| StrSubMany.pdf | User's Guide for the StringSubMany Utility program |
| MathMLFontFix.txt | A String File with binary character subs for UTF-8 MathML |
| HTML2TABLfix.txt | This file was used to convert HTML tables to plain text, but is no longer needed. |

The HTML Table substitution is no longer needed since the 2015 code is completed and all HTML tables have been replaced with properly spaced plain text.

The Utility is now in the **SPECIAL** directory on the **FTP server** as a file named **StrSubMany.7z**. The archive contains our custom Windows command-line based String Substitution utility named **StringSubMany**, along with the User's Guide (**StringSubMany.pdf**), plus the **String Sub text file** named **MathMLFontFix.txt**

As always, **BACKUP YOUR DATA FIRST**, and you assume all risks associated with using software from IP Data Corporation.

Appendix P - FOR THE PROGRAMMER

Useful Tables, Data Sets and Code Samples

We have a few useful data sets and translation tables available **free for subscribers to our data**. These include:

Complete list of *Numeric and Named HTML entities* found in 35 years of patent data, with Unicode Code Points as well as the UTF-8 codes for them. This includes a list of sources and/or references we used to build the tables to trans-code the files into UTF-8.

A translation List of *English Symbols/Words to Unicode Code Point values* (with UTF-8 codes) for the *most common symbols and special characters* we have trans-coded to UTF-8 data. They include symbols and operators for scientific, chemical, mathematics, etc., that can be used to add easy English language searching to a system (as we do in our P9 Search system). This allows someone to easily provide the users of a system familiar terms to use for their technically oriented searches. The table also includes “alternate” (usually less popular) choices we found that patent authors have made in the data (some “OK” and a few selections that left us a bit puzzled and took some time to figure out. Table sample:

| Symbol/Word | Char(s) | Unicode (1) | UTF-8 | Unicode (2) | UTF-8 | Comment |
|-------------------------------------------------------------------------------------------------------------------|---------|-------------|-------------|-------------|-------|---------------------------------------------------|
| Angstrom | Å Å | U+212B | e2 84 ab | U+00C5 | c3 85 | Needed Angstrom, used A-ring |
| Angle | ∠ | U+2220 | e2 88 a0 | | | |
| ... | | | | | | |
| Invisible Times | 3x | U+2062 | e2 81 a2 | | | Yes, it's between the 3 and x (can't you see it?) |
| ... | | | | | | |
| (1) most often used by patent authors/inventors (usually correct), (2) Or this... (usually not correct, but “OK”) | | | | | | |

Utility Programs to Manipulate & Verify MAPS Data

We have a collection of utilities and software “snippets” to assist programmers and system designers who are using the MAPS data. IP Data Corporation offers several command-line based program (utilities) to purchasers of and subscribers to the MAPS data, upon request.

These programs perform functions such as:

- Converting MAPS to other formats like Plain Text, well formed and Valid XML, and the BRS Load format (BRS is also known as the Live Link Discovery Server which is a system used by the USPTO).
- Perform basic integrity checks on MAPS and MAPS-XML data sets for mandatory fields, basic field contents in important fields, properly formatted dates in all date fields.
- Read MAPS files and use the Issue Date or Date Published to the store files in a destination path by Year and date. This is good for storing groups of updated files provided that are from several different publication dates.

We also have code snippets or modules for things like basic display of MAPS data in a CGI-based web server environment, MAPS parsing, complete C Structure array with all known HTML Entities to the UTF-8 code with names/comment and more. **If you need something, ask us!**

APPENDIX S - Search Report Text Data Format

The EPO has standardized and tagged all of the the elements contained in a Search Report including each of the **Citation Types**, it's **Category**, and the **Relevant claims** (by number) and/or **Relevant Text** (by application paragraph number) for the applicable Citations.

We include this in the EP applications that contain the ISR, and we are also working on software to add the same type of labels in PCT Search Reports if they begin to include them again. The initial OCR copies we received were not of sufficient quality to allow us to identify the elements in the PCT copies with a decent measure of reliability. Hopefully this will improve in the future as more electronic sourced data becomes available. The EPO disclaimer advises the user that the Search Report is OCR data, and that it may have errors, and we can tell you that is does have errors, but far fewer than any PCT OCR sourced data.

Two Main Sections in the International Search Reports

Search Reports contain two main sections which are **the Citations**, and **the Family Document Annex** (a list of family documents).

We have converted the EPO Search Report Data to a parsable text format where each of the Citations and its type (e.g., patent, NPL, etc.), the category under which it was cited, and the relevant text or claim (reason it was cited). Cited documents can usually be pulled out of DOCDB (*and they are NOT OCR*), so we recommend that you use DOCDB if you really need an accurate list of cited documents for an application that has had its preliminary search completed. But, it is here if you need it. We index the entire search report as a block of searchable text in our systems. Also, it may not be obvious looking at the included Search Report Text example below, but each of the Citations is on a line by itself after the STXT Maps Paragraph code and the abbreviation **SRep** (indicating this citation was from the Search Report), like this:

STXT SRep Cited: EP0938243A2 19990825, FUJITSU LTD [JP], category: X, ...

Below is a list of all of the tags we have added for the ISR text elements. When the PTO begins to ship their text (hopefully in 2016) we intend to add MAPS Codes and MAPS-XML tags for all of the search report data. We also believe (and hope) the EPO will be shipping electronic versions of the search reports by the end of next year (2016).

The various elements in the Search Report data can be located by searching for the **Prefix Text Labels** we have added, such as "**Cited:**" (or "Cited" plus the NPL type), and finally the "**category:**" followed by the single letter. Below is a list of the Sections, and we have included the labels for each of the elements followed by a colon, and then the data. The element labels include:

| | |
|------------------------|----------------------------------------|
| File Reference ID: | (EPO internal reference number) |
| Application Reference: | (standard EP Applications Number) |
| Report Established: | (Yes or No) |
| Abstract Approval: | (Yes or No) |
| Figure Info: | (Figures submitted with applications) |
| Figure to Publish: | (Representative figure for front page) |
| Search Report office: | (Two Letter Country) |

Date Report mailed: (yyyymmdd)
 Fields Search (min. documentation): (Section, Class and Subclass of the IPC e.g., **H01F**)

The common citation types begin with the following prefixes:

SRep Cited: (for patent documents)
SRep Cited NPL (Serial/Journal/Periodical): (magazines, etc.)
SRep Cited NPL Book, Author: (Published Books)
SRep Cited Online, Author: (Web Site Links)
SRep Cited Article: (Not sure how this differs from Periodicals)

The Citation types are followed by the Category (A Capital letter or &) which are defined on Page 1 of every Search Report. We add the prefix category shown here:

category: (one letter follows **category:** X, Y, A, O, P, T, E, D, L, &)

Following the category, are the Relevant Claims and/or Text passages and possibly a list of applicable figure numbers that are prefixed as shown next:

Relevant claims(s):

Relevant text passages:

Figures:

Below each of the Citation Types listed below (**bold face type**), are indented lists of data element prefixes we use to identify the additional items sometimes used to further describe the various citations. We have found several examples where some of these additional items did not make it from the Report into the text supplied by the EPO. Use with care, and caution.

SRep Cited: (a patent publication number follows "Cited:")
SRep Cited NPL (Serial/Journal/Periodical): (for NPL Serial publications, Journals, Periodicals)
 Imprint: Serial-title: Alt-title: Issue:
 Pub-date: PubID: VID: (Volume ID) Issue-No.:
 DigitalObjectID: ISSN: ISBN: (c): (copyright)
 Notes: AbsNo.:

SRep Cited NPL Book, Author:
 Imprint: Title: Subtitle: Ed.: (edition)
 VID: INO: Notes: PubID:
 Book No.: Class: Keyword(s): (c): [a copyright notice]
 ISSN: ISBN: Refno:

SRep Cited Online, Author:
 Title: Host-title: Ed.: ABSNO: Avail.: DateCited:
 SrchTerms: SrchDate: Refno:

SRep Cited Article:
 Author: Article title: AbsNo.: Class: ArtID:
 (c): Refno:

Example of Search Report Text

The Element Labels are underlined below to help identify them in the text:

SRPT Begin International Search Report

CMNT This text has been acquired by a variety of automated means (e.g., OCR)
 CMNT that may cause inaccuracies. The search report text data is provided by
 CMNT the EPO for the user's convenience only. It might differ from the search
 CMNT report of the PDF [or TIF] document, which contains the officially
 CMNT published data. The EPO disclaims any liability for incorrect or
 CMNT incomplete data in the XML text [or formats derived from XML] for
 CMNT search reports.

STXT File Reference ID: 106 251 a/np0

STXT Application Reference: EP04028357.4

STXT Report Established: yes

STXT Abstract Approval: yes

STXT Figure Info: by-applicant

STXT Figure to Publish: 1

STXT Search Report office: DH

STXT Date Report mailed: 20150107

STXT Fields Search (minimum documentation): H04J

STXT SRep Cited: EP0938243A2 19990825, FUJITSU LTD [JP], category: X, Relevant claims(s): 1-27, Relevant text passages: * paragraph [0001] */* paragraph [0029] - paragraph [0040] */* paragraph [0043] - paragraph [0048]; figure 4 */* paragraph [0077] - paragraph [0097]; Figures: 10,11,12,13,14 *

STXT SRep Cited NPL (Seial/Journal/Periodical): Author: WEI S HU ET AL, Article title: Multicasting Optical Cross Connects Employing Splitter-and-Delivery Switch, Serial-title: IEEE PHOTONICS TECHNOLOGY LETTERS, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, Pub-date: 19980701, VID: 10, Issue-No.: 7, ISSN: 1041-1135, Refno: XP011046132, category: X, Relevant claims(s): 1-27, Relevant text passages: * sections III and IV; Figures: 3,4,5 *

STXT SRep Cited: US6496289B1 20021217, KUROYANAGI SATOSHI [JP] ET AL, category: X, Relevant claims(s): 1-27, Relevant text passages: * column 1, line 7 - line 9 */* column 14, line 32 - column 15, line 14; Figures: 15a,15b *

STXT Primary Examiner: Roldán Andrade, J

STXT Search Report office: The Hague

STXT ===== Begin Patent Family Annex =====

STXT This annex lists the patent family members relating to the documents
 STXT cited in the above Search Report. The members are as contained in the
 STXT European Patent Office EDP file. The European Patent Office is in no
 STXT way liable for these particulars which are for the purpose of
 STXT information only. For more details about this annex:

STXT See Official Journal of the European Patent Office, No 12/82.

STXT SRep Family Priority Appl.: EP0938243A2 19990825

STXT Family Member: DE69833591T2 20070201

STXT Family Member: DE69838353T2 20080521

STXT Family Member: EP1467591A1 20041013

STXT Family Member: EP1467592A1 20041013

STXT Family Member: JP3425861B2 20030714

STXT End of Family Members

STXT SRep Family Priority Appl.: US6496289B1 20021217

STXT Family Member: JP3442277B2 20030902

STXT Family Member: JPH11239368A 19990831

STXT Family Member: US2003063347A1 20030403

STXT End of Family Members

As always, we highly recommend that you always check the facsimile copy (the legal copy) for any legal purposes or requirements.

Appendix T – ST.8 Standard Classifications

Details on USPC, CPC and IPC Class formats

The USPTO goal of designating the US Classification System *STATIC* appears to be a reality (except for Design Patents, of course). If you are not familiar with the CPC, you should probably put in some overtime and get caught up on it.

The EPO and the USPTO have worked for two solid years to make the CPC a reality, and “we have arrived.”

All through 2013, the CPC took a back-seat to the US Class system while the technology centers got up to speed, and a month or two into 2014, they threw the switch and the CPC jumped into the front seat and took over as the primary US classification system.

The EPO had embraced it 100% from the start, but it was much easier for them since the CPC was a direct conversion from ECLA, now a “dead” system and a bumpy foot note in the history of patent data. **It was not so easy for the USPTO.**

CPC and IPC Both Use ST.8 Standard

With the CPC and IPC both using the ST.8 standard, many users can standardize their processing, storage and interpretation of the IPC and CPC entries. For example, we are changing the way we normalize IPC classes internally in our systems to match the new **CPC normalization**. IPC versions 1 through 8 never used more than 3 of the Main Group digits, whereas the CPC uses all 4, quite often. When we parse and store the IPC, we now pad the Main Group to store and index all 4 digits. This way, the same processing functions used for sorting, storage, indexing, and even for display will work for both IPC and CPC entries. We think it is a nice “bonus” for switching to the CPC!

Only 15 of the 50 Bytes Define the Technology Class

Almost all new IPC and CPC class entries have 50 characters, even though it only takes 15 characters for the entire Classification portion of the entry.

Display and Storage of the 15 Class Characters

For example, here are the two entries (IPC and CPC) from **US5590424**:

In the next examples, the Unicode “**Visible Space**” character is shown in place of spaces:

CPCP A47K__13/165 (20130101)

A47K0013/165__

Normalized for storage, sorting, display and manipulation on some systems (ours in particular)

CLSI A47K__13/14__

Shown for display in some systems (spaces vs. zeros)

A47K__13/14__

Normalized spacing

CLSI A47K_13/14

Older IPC Display Format - Many of the older IPC entries were displayed with three characters in the Main Group (before /) since older IPC versions never had more than 3 digits (we think we have them all updated to 4 in of the new data sets).

Zero Filling the main group allows proper sorting of the entire 15 byte entry.

The first **normalized** entry shown with zero-filled Main Group and space-padded trailing Subgroup characters allows certain numeric sorting techniques and DBS index setting that may cause problems with spaces. Only your programmers will know for sure.

IMPORTANT: You must NEVER add zeros to the end of the SubGroup since this will change the entry to a different legal Class value and will prevent you from recovering the original value since both are legal entries (this is mentioned a few more times due to its importance, not because we forgot we already mentioned it).

This table shows the lengths and positions of each of the character fields discussed (and maybe one or two are not).

Class Data Format Table from ST.8 WIPO Standard

| Position(s) | Content | Values |
|--------------------|-----------------------------------|----------------------------------------|
| 1 | Section | A, ..., H |
| 2,3 | Class | 01, ..., 99 |
| 4 | Subclass | A, ..., Z |
| 5 to 8 | Main Group (right aligned) | 1, ..., 9999, blank (note 1) |
| 9 | Separating character | / ("Slash") |
| 10 to 15 | Subgroup (left aligned) | 00, ..., 999999, blank (note 2) |
| 16 to 19 | For future use | 4 blanks |
| 20 to 27 | Version indicator | YYYYMMDD date format |
| 28 | Classification level | C, A, S (note 3) |
| 29 | First or Later position of symbol | F, L |
| 30 | Classification value | I, N (invention or additional) |
| 31 to 38 | Action date | YYYYMMDD date format |
| 39 | Original or reclassified data | B, R, V, D |
| 40 | Source of classification data | H, M, G |
| 41-42 | Generating office | AA, ..., ZZ (per ST.3) |
| 43-50 | For future use | 8 blanks |

Unused positions in the IPC classification fields Main Group (positions 5-8) and Subgroup (positions 10-15) should be left blank for display (preserving spaces). The only other positions that may be left "blank" are the ones reserved for "future use". All other positions must be assigned one of the acceptable "values" listed in the table of paragraph 3 of the specification (not included here). Any zero appearing in the symbols should be recorded (e.g., used, not removed)

NOTE 1: 1 to 4 digits. Right-aligned to the slant bar (9th position). **Leading zeros are NOT significant and ST.8 recommends they be left blank**, but you may find them filled on some systems where they are padded for sorting or storage purposes. The field may also be left Blank (no data).

NOTE 2: 2 to 6 digits. **Leading and trailing zeros ARE significant** (leave this entry exactly as delivered). This field may also be left Blank (no data).

NOTE 3: The **Classification Level is not used in the CPC entries**. They have however added another single character field called **scheme-origination-code** (which is mistakenly labeled scheme-organization-code in some of the USPTO AIA-CPC documentation). This is a single digit and describes the source of the CPC entry design or scheme (we don't have all of the details, yet).

There are a total of 50 characters (bytes) defined in ST.8, but for indexing and search purposes most systems only index the classification itself (15 character, shown above). The additional fields are useful for many things, but not for searching by classification except where systems have separate indexes for Primary class, Invention-Related and Non-Invention-Related classes, and CPC Combination Classes.

Both the IPC and CPC are stored with spaces between the CLASS and GROUP section. MAPS stores them as received in the subscriptions we use to build the file (all 50 characters including reserved spaces).

If you are a stickler for precise data formats and you want CPC and IPC classes to sort correctly, you can pad the IPC to 15 full characters as shown above **with leading zeros on the Group numbers**, and preserving the trailing spaces. The above example would look like this: **A06L0002/13** (with 4 trailing blanks) – **and, you should NEVER zero pad the trailing spaces** because /13 is not the same as /130 and is not the same as /1300 and all three are different, probably unrelated group/subgroup values.

The CICE MAPS Field indicates which “edition” of the IPC the classes were assigned under. The edition listed in this field applies to the IPC classifications listed directly after the entry. This code is not used but remains documented since it is still in the Authority specifications. The IPC and CPC classes now use the Version Date in the 50 character class entry as the version indicator.

Appendix U – Unicode UTF-8 Character Sets for MAPS

UTF-8 (8-bit UCS/Unicode Transformation Format) is a variable-length character encoding for Unicode created by Ken Thompson and Rob Pike.

UTF-8 is able to represent any universal character in the Unicode standard, yet the initial encoding of byte codes and character assignments for UTF-8 is coincident with ASCII requiring little or no change for software that handles ASCII but preserves other values. For these reasons, it is steadily becoming the preferred encoding for email, web pages, and other places where characters are stored or streamed.

UTF-8 uses one to four bytes (octets) per character, depending on the Unicode symbol. **Single Byte Codes** encode the 128 US-ASCII characters (Unicode range U+0000 to U+007F) and the ASCII values (0 to 127) remain literally unchanged. **Two Byte Codes** are used to encode most Latin letters with diacritics and for characters from Greek, Cyrillic, Armenian, Hebrew, Arabic, Syriac and Thaana alphabets (Unicode range U+0080 to U+07FF). **Three Byte Codes** are used for the rest of the Basic Multilingual Plane which contains virtually all characters in common use around the world. **Four Byte Codes** are used to encode characters in other, lesser used, planes of Unicode.

UTF-16 (the main alternative to UTF-8) also needs four bytes for these code points. Whether UTF-8 or UTF-16 is more efficient depends on the range of code points being used. However, the differences in size between different encoding schemes can become negligible with the use of traditional compression systems like DEFLATE.

All recent version browsers will recognize and display UTF-8 data properly when set to “Automatic” for character set recognition, with the exception of Math-ML data. Math-ML Data is discussed on **Appendix M**, titled **Math-ML data**.

Supported UTF-8 Ranges (all Languages)

In addition to the 127 character single byte ASCII set, we have identified the following Unicode Ranges as the complete set of Unicode Characters used in all weekly published patent documents from the four Authorities we currently support:

| Description | Unicode Range (hexadecimal) | Our Header file |
|--------------------------------------------------------------------------------------|------------------------------------|------------------------|
| (CJK is Chinese, Japanese, Korean) | | |
| English, European (Latin-based), Greek, Hebrew | 1– 2351 (x0001 - x092F) | “UTF-8-Basic” |
| Math Symbols, Special Symbols , Letter-like symbols, drawing characters, etc. | 7424 –11097 (x1D00 - x2B59) | “UTF8-MathSym.h” |
| CJK - Symbols and Punctuation | 12288 – 13311 (x3000 - x33FF) | “UTF8-CJK-Sym.h” |
| CJK - Unified Ideographs - Extended Set-A | 13312 – 19894 (x3400 - x4DB6) | “UTF8-CJK-ExtA.h” |
| CJK - Unified Ideographs - Main Set (20,939 ideographic characters) | 19968 – 40907 (x4E00 – x9FCB) | UTF8-CJK-Main.h |
| Korean Hangul Syllable Ideographs - Main Set | 44032 – 55291 (xAC00 – Xd7fb) | “UTF8-Hangul.h” |

| | | |
|----------------------------------------------------------------------|------------------------------------------------------------------|------------------------|
| Math-ML Markup Symbols - Unicode Private Area | 57344 – 63743 (xE000 - xF8FF) | UTF8-Private-MathML.h |
| Math-ML-2 Markup Symbols - Unicode Private Area 2 | 63744 - 65279 (xF900 – xFEFF) | “UTF8-Private-Math2.h” |
| CJK - Wide Punctuation , Form Characters and Small Characters | 65280 - 65535 (xFF00 – xFFFF) | “UTF8-Wide-Punct.h” |
| MathML MarkUp Symbols in Unicode Private area 3 | 119808-120831 (x1D400-x1D7FF) (1024 characters above U+0FFFF) | “UTF8-MathML-4byte.h” |

IMPORTANT NOTE: Do not edit MAPS text files unless you are absolutely sure that your editor fully supports UTF-8 character encoding.

For Windows based computers, we recommend the **Notepad2-mod text editor**, originally written by Florian Balmer (a high quality “Cracker Jack” programmer from Switzerland) and now “beefed up” and extended by the Open Source community (hence the “mod” suffix).

Here is why we recommend it:

- ⤴ We have come to trust Florian's code and the newer Group's code over the years.
- ⤴ Fully supports UTF-8 and other character encoding schemes, plus it allows you to manually recode the text forcing UTF-8 even when a code is improperly encoded, and it always correctly FLAGS the bad conversion characters.
- ⤴ It converts from ANSI to UTF-8 and back (and to and from any supported sets)
- ⤴ Syntax Highlighting for multiple programming languages,
- ⤴ As long as lines are not overly long (< 4K), it opens very LARGE line oriented files (~100 megabytes in 10 seconds).
- ⤴ You do not need to install it (a Portable Version is available),
- ⤴ It does not modify your Registry at all.
- ⤴ Works on Windows 2000 through Win 10 and all server versions in between.
- ⤴ It's FREE (Source code is available on GitHub.com)
- ⤴ Now with a Super Improved version with Code Contributors building on Florian's work!

The original is here: <http://www.flos-freeware.ch/notepad2.html> and

The “Beefed up” Notepad2-mod is here: <http://xhmikosr.github.com/notepad2-mod/>

A Portable Version is located here: http://portableapps.com/apps/development/notepad2-mod_portable

For more editing power that handles UTF-8 correctly (such as inserting columnar data, more Macro power, extensible, multi-windowed, and much, much more) **AkelPad** is another option that we use. More information is available here: <http://akelpad.sourceforge.net/en/index.php>

Indexing UTF-8 Data with Symbols and

Most modern DBS platforms handle most of the character parsing, indexing and searching details dealing with various character sets, but if you are designing your own methods and algorithms to handle the special characters and binary data for data entry entry, search comparisons, special display methods, there are two ways that we have used to approach the design:

1. Your parser can index the symbols as text by their name ("iota", "omega", "times", etc.) which means

you can probably use ASCII text for everything except large Ideographic character sets (e.g., CJK, etc.) if you index them, at all.

2. Index them as UTF-8 binary or even as Unicode Code Points and create a way for users to insert these characters into their query. Then you can accept text names and then search as UTF-8 or Unicode Code Points (whichever way you indexed them).

Both ways require look up tables of some sort, and as mentioned, we have tables for our subscribers for all of the symbols we have run across and converted for our conversion project.

HTML Entities of various kinds were included by the Patent Authorities, including:

- Named Entities, for example: ′ &sqr; &ga; &Ga; etc.
- Decimal Numeric Entities: {
- Hexadecimal Numeric Entities: †

Most of them are in the **.ent** files that come with the Schema (or DTD package).

In the early years of SGML, HTML XML, XHTML etc., some really STRANGE "Ampersand Entities" kept showing up from the different patent authorities, and at one point we were absolutely certain that these codes were plucked directly from Alien-Space-Craft passing by at near light speeds, since no one seemed to be able to find the source for them. This was especially a problem with WIPO documents since they were OCR sourced. Suffice it to say, it was not a straightforward process converting all of the data from multiple Authorities into a single binary data set, and many of complications arose from changes in the sets that happened every few years over the 30 to 40 or so years that most of these collections cover.

Many of the errors are handled with software, but the number of hours we have spent editing files is phenomenal and feels like we are in competition with the labor hours expended on the Grand Coulee Damn (and, at \$4.9 Million for labor on the damn, we may actually reach that some day – JUST EDITING...).

APPENDIX W – WIPO Publication Number Details

Introduction

This appendix discusses the PCT Publication Numbers including format changes over the years, details on Kind Codes for corrections and revisions, and finally, how IP Data Corporation (IPDC) filenames are created from these numbers to store and track the PCT World Application text and image files. We also discuss a major change in PCT Publication Number formats as we follow the EPO and DOCDB and **not WIPO**, and how this affects publication numbers and filenames. First, we discuss how the basic PCT numbering has morphed over the years.

PCT Publication Number Formats from 1978 to Now

When WIPO began publishing patent applications under the Patent Cooperation Treaty in 1978, they did not foresee the popularity of the Treaty. We know that computer systems were far less powerful back then, and saving every byte was important, but from the standpoint of a computer programmer they really made some serious mistakes in the format of publication numbering for the PCT applications. It may have been due to bad program managers, poor computer software engineers, or just a simple lack of foresight. But, whatever the reason, it led to the confusing mess of publication numbers that it takes a computer to figure out and keep straight.

IMPORTANT NOTE: WIPO's systems do not recognize DOCDB Kind Codes for Pre-2009 WO Publications. The EPO has assigned older Corrections and Amended Claims documents Kind Codes described in the WIPO Standards which are the same ones they use with EP Publications (except for the Kind A4 which are Amended Claims on WOs and a special search report document for EPs that they do not distribute with the weekly data or the back-file collections).

IP Data Corporation publication numbers follow the EPO's format for PCT publications which include the Kind Codes from DOCDB prior to 2009. WIPO does not yet recognize these Kind Codes (but they should).

Initial PCT Publications Number Format in 1978

PCT applications are created and submitted all over the world. But the numbers they are assigned are handed out by WIPO (the World Intellectual Property Organization). In 1978 they were fairly far from the turn of the century, and no one working then looked far enough ahead to see what could happen.

For the newly published applications they used a 2-digit year, that was the last two digits of the Gregorian calendar, followed by a 5-digit sequence, then followed by the document Kind Code (A1, A2, A3, etc.).

W07800001A1 (World, 1978, 00001, Kind A1)

The 5-digit sequence was fairly short lived. But you can even find some PCT number with a 4 digit sequence.

PCT Publication Sequence Number Goes to 5 Digits in 2002

The 5-digit sequence number lasted all the way to June 26th, 2002, when the last one was published. It was number sequence number 51230. At this point they saw the writing on the wall – they were going over 5 digits before the end of 2002, like it or not! All applications the following week (July 4th 2002) used 6 digits. This gave us:

W002051231A1 (World, 2002, 051231, Kind A1)

In fact, they published 104091 applications in 2002, the last one being WO02104091A1.

PCT Publication Number “Modernized” in 2004

In 2004, someone looked ahead and said, “Hey, this patent thing is catching on!” and they could see that in 2078 there could be some real trouble. So they finally bit the bullet and went to a 4 digit years. This now gives us the current WIPO number format of a 4 digit year and 6 digit sequence, like this:

W02004000001A1 (World, 2004, 000001, Kind A1)

The US has been doing patents for a while, and when they began publishing the applications in 2001, they went with a 4-digit year and a 7-digit sequence number. But do you really think there will be more than 999,999 US applications published in one year? It is possible when you consider there were over 300,000 US applications published in 2007 and we are now very close to 400,000! But, it should be a few years away.

Something to note (for you programmers out there) is that when corrections or A3 documents publish for the older numbers, they do not use the old number format. They use the latest format (4-6). So it is possible to have a set of related publications with two or even 3 formats in some databases around the world:

WOYY12345A2 then the A3 as **WOYY012345A3**, then a correction as **WOYYYY012345A3**

IP Data Corporation now stores and handles all PCT publication numbers including filenames using the 4-digit year and a 6-digit sequence. This include the list and references inside the PCT documents.

EPO Accepts Some PCT Applications as Native EP Applications

Anyone working with WIPO and EPO patent data is probably aware that over the 35 years since the **Patent Cooperation Treaty (PCT)** was implemented, many patent authorities have accepted and used the PCT applications, but not always in the same way.

EuroPCT Applications

The European Union accepts WO applications as Native EP applications, **but only if the application meets certain specific criteria** such as being submitted in one of the 3 approved languages. When they accept the WO as the Native application, this is known as a **EuroPCT**

The **PCT** and the **EPC** have matured together. Prior to 2010, the EPO handled all of the WO application data for WIPO, including databases and bulk data distribution. In 2010 WIPO assumed the responsibility to maintain and distribute all of the PCT (**WO**) Application data.

Prior to approximately 1990, less than 10% to 15% of the PCT applications submitted to the EPO were accepted as Native **EP** Applications, but over the years it has grown, and by the end of 2014, it was typically 30% to 35% each week. **This means there are no EP applications for approximately one third of the numeric series of EP Application numbers.** These applications are called **Euro-PCT Applications**. An EP “stub” (our name for them) with a reserved EP sequence number is included in the weekly data with basic bibliographic information and the PCT Number so that you can easily locate the complete PCT application that was submitted. If the application is granted, it is published as a **Kind B** EP document with that reserved sequence number.

Not all WO applications are accepted by the EPO as Native EP applications. In these cases there is a parallel EP application and both exist. Lists of all Cross-References where the WO is accepted and no EP exists, plus lists where both exist can be found here for years 2005 through the current issue week:

<http://www.epo.org/searching/free/publication-server/cross-references.html>

Also, a link to lists of unpublished EP numbers that are either withdrawn or unused for various reasons, can also be found on the same page.

IP Data Corporation has compiled the list of Euro-PCT applications prior to 2005 and it is available upon request.

Prior to 2009 WIPO Does Not Use Correction Kind Codes

Prior to 2009, WIPO did not publish Corrected Copies of their documents using the A4, A8 and A9 Kind Codes listed in their recommended standards. The **Kind Codes on the Corrected and Revised copies were the same Kind Codes found on the original publication.** They simply added a stamp across the top saying “Revised” or “Corrected” and the date of the corrected copy was listed in **INID section (48)** along with the **Correction History which was listed in INID Section (15)** which lists the date and PCT Gazette number of the current and all prior corrections. These two INID entries were most often found on page 2 of the facsimile image document at the very end of the Bibliographic data (in small print). Also, **Amended Claims** usually did not have the Corrected or Revised Stamp at the top of the image copy, and the only indicator that it was a correction was a statement describing that it was Amended Claims in **INID Section (88)**, also usually on page 2 (in small print). But it was usually obvious that it was Amended Claims if you looked at the first page past the bibliographic data because the first herald at the top of that page was “Amended Claims” along with a date the claims were received by “the Bureau” which is another name for World Intellectual Property Office.

When you created a filename with the Country code, the Sequence and the Kind Code supplied in the electronic copies, you ended up with the same publication number as the original document. We wanted **unique filenames** for all PCT Text and Image files and we initially added a simple suffix of -1, -2, -3 etc. to the corrections keeping the Kind Codes WIPO left on them. This lasted for several years until the EPO took the initiative and assigned the proper Kind Codes to the Pre-2009 PCT Corrections and documented in DOADB.

EPO Renames WO (PCT) Correction Documents (new Kind Codes)

In 2012, the EPO renamed all of the Pre-2009 PCT publications and added them to DOCDB with the new numbers. Since we use DOCDB, we followed the EPO and in December of 2014 we renamed all of the Pre-2009 publications to match DOCDB.

Additional Details on PCT (WO) Kind Codes

PCT (WO) Primary Documents

PCT Applications initially publish as either **A1** or **A2 Kind** documents. **A1 Kind documents** indicate that the preliminary **International Search was completed** and published **with** the application. Approximately **2,002,194** Kind A1 applications were published from 1978 through 2014.

The **A2 Kind Code** indicates that the **search report was not ready at publishing time** and the ISR would be published separately at a later date as a **Kind A3** document. Approximately **560,360 PCT A2 Kind** applications were published through the end of 2014.

The Search Report is included with the Facsimile Image copy (usually 2 to 5 printed pages), but the **Search Report Text is not currently included with the PCT A3 Text file** shipped from WIPO.

When we purchased the “new” Back-file rebuilt by WIPO in 2011, there were no A3 publications included with it prior to 2002. We created approximately 15,000 missing A3 publications from the DOCDB data.

Correction Documents

Along with the three initial document Kinds listed above (A1, A2 and A3), there are also correction documents which are corrected copies or partial “Front Page” corrections published at some point after the initial, primary publication. These are often referred to as “**Further Publications.**” The most common Kind Codes across the industry for these publications are currently **A4, A8** and **A9** described next:

PCT Amended Claims Use Kind Code A4 (or B1 prior to 2009)

Kind A4 publications contain updated bibliographic data (front page) and Amended Claims text, only. Prior to 2009, **Kind B1** is used for PCT Amended Claims document. We have no idea why they are Kind B1, plus, the B1 Amended Claims documents were not provided in the new PCT back-file in 2011. We created the B1 Kind documents with the bibliographic elements from the DOCDB data.

IMPORTANT NOTE: Please note that the Amended Claims in PCT A4 publications do not always replace all prior claims text. If you maintain a database of text independent of the documents, **you should always ADD (append) the Amended Claims to the existing claims in your DBS** to be sure you do not replace or overwrite valid claims that were not amended. Normally, the word [Amended] or [Deleted] is supposed to precede amended or deleted claims in the PCT A4 Amended Claims documents, but not all applicants follow the recommendations.

Kind A8 publications correct Front Page data also called Bibliographic Data. The description of A8 Kind documents states that all data (each of the individual items) on the correction is the latest data available and that each A8 will contain all prior corrections, which means they are **CUMULATIVE**.

Kind A9 publications are a bit more complicated in PCT documents. The A9 is a complete replacement of a publication or an addition to an existing publication such as a translation. What complicates the A9 Kind is that the publication it replaces can be the Primary document (e.g., the Full Application), or it may be the Search Report, or it may be adding a translation of an existing search report. This means that if you are updating a database and you want to provide your users with the latest publication, **you must be sure you check the correction codes to be sure you know what document the A9 is replacing.**

The relationships of Correction Codes and Kind Codes are discussed in more detail in:
APPENDIX C, titled Correction Codes for PCT (WO) Pubs

NOTE: If you receive a set of images from IPDC (TIF or PDF), the Images may still use the -1, -2, -3 , etc., numbering scheme on Correction Documents until we rename the images in late 2015 or Early 2016. We will replace all renumbered Correction images to any customer with the PCT Image back-file. The delay is caused by the fact that we are modifying the faces of the documents to add the new Publication Number and Kind Codes across the top of the front page of each file.

Appendix Z - Abbreviated Table of MAPS Codes

Each Group or Section will have every MAPS code listed that may be used in that Group or section except CMNT (comment) which may appear in any section.

| MAPS | Used In | INID | Name/Description |
|--------------------------------------------------------------------------------------------------------|----------|------|----------------------------------------------------------------------------------------|
| Publication Information Group [10] – Identifies the Current Publication | | | |
| PATA | USA | [11] | Identifies US Applications with Publication Number |
| PATB | USB | [11] | Identifies US Granted Patent Publication Number |
| USAA | USAA | [11] | Identifies US Application Publication Number in Abstract |
| USPA | USBA | [11] | Identifies US Granted Patent Publication Number in Abstract |
| PATW | WO | [11] | Identifies WO Publication Number |
| PATU | EPA EPB | [11] | Identifies EP Publication Number |
| PAJA | JPAA | [11] | Identifies Publication Number for JP Abstract. |
| PNAL | ALL | [11] | Publication Number Alias – Optional lists older Pub Number format, if still valid |
| ISSD | USB, EPB | [45] | Issue Date for a Patent. Same purpose as PBDT, except it indicates a Grant is Issued |
| PBDT | ALL | [14] | Date Published – for current document. Format: YYYYMMDD |
| APPD | EP WO | [43] | Date of Publication of First Application (usually the A1 or A2 Doc Kinds) |
| PKCD | ALL | [13] | Publication Kind Code – 1 or 2 characters. Usually 1 letter, or 1 letter plus 1 digit. |
| PBLN | ALL | [26] | Publishing Language – Language in Desc and Claims |
| LABL | ALL | [12] | Document Name Label - 1 or 2 optional lines add to description or purpose of pub |
| MVER | ALL | | MAPS Version and MAPS File Creation Date |
| Correction Information Section [15] (not a Group since it is a part of the document ID) | | | |
| CORR | EP WO | [15] | Correction Information - Type of correction, Bulletin number, Affected Section |
| CODT | EP WO | [48] | Date of Correction – Applies to CORR line above it, YYYYMMDD Bulletin:YYYYNN |
| DSRP | EP WO | [88] | Date of Search Report – Date of this or earlier Search Report. May follow CORR |
| Application Information Group [20] – Identifies Original Filing Information | | | |
| APNM | US | [21] | Application Number - (US, 6 digit) – SCOD plus this APN forms unique US APNO |
| APNO | EP WO JP | [21] | Original Application Number - Assigned by the authority (EP, WO or JP). |
| SCOD | US | [21] | Series Code – 2 digits 01-14, Design is letter “D” or 29. SCOD+APNM = NN/123456 |
| APDT | ALL | [22] | Application Filing Date – YYYYMMDD - follows APNM or APNO |
| AFLN | ALL | [25] | Filing Language – of Original Application. 2 lower case letters from ISO-639. |
| PRLN | EP | [27] | Processing or Procedure Language - Language used to process the application. |
| FMID | ALL | [30] | Family ID – DOCDB integer defines apps and patents sharing common Priority Claim. |
| Priority Application Listing Group [30] | | | |
| FADL | ALL | [30] | Priority Application Listing – Date Filed, Country, Application Number |
| Public Availability Dates Group [40] - Dates of various phases of Application to Grant and more | | | |

| | | | |
|---------------------------------------------------------------------------------------------------------|-------|------|----------------------------------------------------------------------------------------------------|
| AMDT | EPB | [45] | Date of Grant and Mention of the Opposition Decision |
| APPD | EP WO | [43] | Date of 1st Published Application – Also used in Pub Info Group in Further Pubs. |
| INTG | EPB | [45] | Date of Announcement of Intention to Grant Without Preceding Exam Report |
| NOIG | EPB | [45] | Notice of Intention to Grant Date - YYYYMMDD (Date only) |
| B241 | EP WO | [24] | Examination Request Date |
| B242 | EP WO | [24] | Dispatch Date of 1st Exam. Report |
| B243 | EP | [24] | Date of Patent Maintained as Amended |
| B244 | EP WO | [24] | Date Request for Conversion to National Application - Contains Date and Country |
| B245 | EP | [24] | Date of suspension (Rule 13) / interruption (Rule 90) of proceedings |
| B246 | EP | [24] | Date of Resumption of Proceedings |
| B248 | EP | [24] | Date of Notification Rights after Appeal |
| B405 | EP WO | [40] | Date and Bulletin Number announcing first availability of this publication. |
| B450 | EP | [45] | Date published and mention of grant (EP B1) or Mention of grant (EP B2) |
| B453 | EP | [45] | Date of publication and mention of the limitation decision |
| B473 | US EP | [47] | Disclaimer Date - Standard Date (YYYYMMDD) plus Bulletin No. in EPB. |
| B47D | USB | [47] | This patent is subject to a terminal disclaimer ← text is always included as data. |
| B474 | US EP | [47] | Term of Grant - Rarely appears in primary EPB docs. Mostly 14 Yrs in US Designs. |
| B47T | USB | [47] | Patent Term Extension – Is “5 years” or an integer 1 to 4 digits for number of days. |
| B475 | EPB | [47] | Lapse of Patent - Date and Country. May have multiple entries |
| Technical Information Group [50] – Various Technical Data Elements in this Application or Patent | | | |
| TITL | ALL | [54] | Title in <i>English</i> |
| TITF | EP WO | [54] | Title in <i>French</i> . EP Grants can contain all 3 EPC languages (en, de, fr) |
| TITD | EP | [54] | Title in <i>German</i> . EP Grants can contain all 3 EPC languages (en, de, fr) |
| TITO | WO | [54] | Title using language in PBLN. (over 30 langs supported by WIPO for PCT). |
| CPCV | ALL | [52] | CPC Version – Used if CPC edition date not listed. Format: yyyyymmdd |
| CPCP | ALL | [52] | CPC Primary Class - Follows ST.8 Standard, mostly |
| CPCI | ALL | [52] | Class, Invention related - second or subsequent classification assigned. |
| CPCN | ALL | [52] | Class, Non-Invention related (or in addition to the invention) |
| CPCC | ALL | [52] | CPC Combination Classes – up to 100 ST.8 Class w/ Group and Rank No. |
| CLSU | US | [52] | US Class/Subclass, Primary entry - (now for Design Patents ONLY) |
| CRCU | USB | [52] | US Class/Subclass, Cross-Reference - (now for Design Patents ONLY) |
| CICE | ALL | [51] | International Class Edition – a digit 2 to 7 |
| CLSI | ALL | [51] | International Class entry (Primary) |
| CCRI | ALL | [51] | Additional IPC Classifications - (Further) |
| FOSS | USB | [58] | Field of Search – Technology fields by Class |
| REFP | ALL | [56] | Cited Patent Reference - a patent reference from any authority. |

| | | | |
|--------------------------------------------------------------------------------------------------|------------|------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| OREF | US, EP | [56] | Cited NPL Reference - NPL Refs incl. books, articles, papers, URLs, etc. |
| Next Six Codes in Technical Group Being Phased Out over 2016 | | | |
| REFS | US EP | [56] | References Cited – Section Heading, will be eliminated over 2016 |
| URFS | USB | [56] | U.S. References – Section Heading, will be eliminated over 2016 |
| FRFS | USB | [56] | Foreign References – Section Heading, will be eliminated over 2016 |
| UREF | USB | [56] | US Patent Ref – in US Pubs only – Phased out 2016 – REFP replaces it. |
| FREF | USB | [56] | Non-US (foreign) Patent Reference – Phased out 2016 – REFP replaces it. |
| ORFS | US EP | [56] | Other References – Section Heading, will be eliminated over 2016 |
| Related Documents Group [60] – National applications and patents related to this document | | | |
| RELA | US EP JP | [60] | Related Application - |
| RELP | ALL | [65] | Prior Publication Data – Pubnum and date of earlier version of this document |
| REIS | USB | [64] | Reissue Data – found in US Reissue Patents prior to Year 2001, ONLY |
| Concerned Parties Group [70] | | | |
| APLX | ALL | [71] | Applicant - 5 sub-fields, Name, addr, city-st, ctry, post-code |
| APPL | ALL | [71] | Applicant – Single text line, loose formatting |
| INVX | ALL | [72] | Inventor - 5 sub-fields provide granularity - See Appendix A for details. |
| INVS | ALL | [72] | Inventor – Single text line, loose formatting, use INVX if possible |
| ASGX | ALL | [73] | Assignee - 5 sub-fields, Name, addr, city-st, ctry, post-code, Role |
| ASSS | ALL | [73] | Assignee – Single text line, loose formatting – Use ASGX if at all possible |
| AGTX | ALL | [74] | Agent - 5 sub-fields provide granularity - See Appendix A for details. |
| AAFM | ALL | [74] | Agent – Single text line, loose formatting, use INVX if possible |
| OPON | EPB | [78] | Opponent and Agent |
| PEXM | USB | | Primary Examiner |
| AEXM | USB | | Assistant Examiner |
| PCT and National Phase Group [80] | | | |
| DSTS | EP WO | [81] | Designated States |
| DSTR | WO | [84] | Designated Regional States |
| IPNO | US, EP, JP | [87] | International Publication Number or Hague Agreement Registration Number |
| PPDT | USB, EP | [87] | PCT Publication Date – or Hague Design Registration published |
| PCNO | EP, US | [86] | PCT Application Number – Different format for US and EP |
| PCFD | EP, US | [86] | PCT File Date or Hague Design File Date filed with the applicable Authority |
| P102 | USB | [86] | PCT 102(e) Date - from 35 USC Section 102(e) before 2003, or Hague Agreement Registration Date 2016 and later in US Design Patents |
| P371 | USB | [86] | PCT 371 Date - refers to date in 35 USC section 371(c)(1), (2) and (4) |
| P372 | USB | [86] | PCT 372 Date - (2015), refers to date in 35 USC section 371(c)(1) & (2) |
| DSRP | EP, WO | [88] | Date of Search Report (w/opt. Bulletin Num) – also used in CORR sect for WOs |
| Abstract Text Section (last item in Bibliographic Data) | | | |

| | | | |
|-----------------------------------------------------------------------------------------------------------|-------|------|-------------------------------------------------------------------------------------------------------|
| ABST | ALL | [57] | Abstract text paragraph in English (en). |
| ABSD | EP | [57] | Abstract text in German (de) |
| ABSF | EP WO | [57] | Abstract text in French (fr). |
| ABSO | WO | [57] | Abstract text paragraph in language listed in PBLN |
| ATXT | ALL | [57] | Abstract Text Paragraph (2 nd or later para. Language same as ABSx above it) |
| TABL | ALL | [57] | Table Line |
| Begin Description Text Section of Application or Patent (primary text of the patent specification) | | | |
| DESC | ALL | | Begin Description in Language Specified in Code PBLN |
| DESF | EP | | Begin Description in French |
| DESD | EP | | Begin Description in German |
| HDR1 | ALL | | Header Text (Level 1) – currently HDR1 is the only one in use. |
| HDR2 | ALL | | Header Text (Level 2) – Not currently used, will be phased in late 2016 |
| HDR3 | ALL | | Header Text (Level 3) – Not currently used, will be phased in late 2016 |
| DTXT | ALL | | Text Paragraph - in Description section |
| TABL | ALL | | Table Line |
| Begin Claims Text Section of Application or Patent (Claims of the patent specification) | | | |
| CLMS | ALL | | Begin Claims Section |
| CLMD | EP | | Begin Claims Section- in the German language |
| CLMF | EP | | Begin Claims Section- in the French language |
| CLMA | WO | | Begin Amended Claims Section – w/ Original Claims in A1 A2, |
| HDR1 | ALL | | Heading text – Used for Inventor's Response in Amended Claims |
| CNTX | ALL | | Numbered Claims Paragraph (whole claim or elements follow) |
| CTXT | ALL | | Claims Text Paragraph Element (added to Numbered claims) |
| DCLM | USB | | Begin Design Claims Section |
| DCTX | USB | | Design Claim Text - Only one unnumbered claim in US Design Patents |
| Begin Search Report Text Section of Applications (A1, A3, A9 or Pre-2009 WO B1 Kind docs, Only) | | | |
| SRPT | EP WO | | Begin International Search Report (←this text is on the line with SRPT) |
| STXT | EP WO | | Search Report Text Paragraph |
| DFAM | ALL | | Family Document |
| TABL | ALL | | One line of table Text data |
| HDR1 | ALL | | Text Heading – Level (1) Primary section names |
| HDR2 | ALL | | Text Heading – Level (2) Subsection names (Example 1, Example 2, etc.) |
| HDR3 | ALL | | Text Heading – Level (3) Used for main patent specific sections |
| Codes optionally used in all sections, or once in all documents | | | |
| CMNT | ALL | | MAPS Comment Field – Used optionally in all document sections |
| PATE | ALL | | Patent End (End of Document) – Must be last MAPS code in all MAPS files |

The **5 Asterisks** have appeared at the bottom of the US Printed Patents for many years to indicate the end of the text. After all, you needed to be sure your last page wasn't eaten by some monster copying machine at the USPTO when you headed out to study 15 pounds of patents copied from "**The Stacks**" (if "The Stacks" doesn't mean anything to you, it's not important – it's just a memory of a different world, now long gone...).

* * * * *