WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

E

PCT International Cooperation Division

DATA FORMAT SPECIFICATIONS FOR THE COLLECTION OF PCT NATIONAL PHASE INFORMATION

Version number 4.0

INTRODUCTION

The purpose of this document is to describe the technical requirements for provision of PCT national phase entry event information "National Phase Entry Information" by Intellectual Property (IP) Offices to the International Bureau (IB) of WIPO.

Following PCT Notification No. 207, relating to the adoption of amendments to the Regulations under the Patent Cooperation Treaty (PCT) adopted on October 14, 2015, the amendments to Rule 95.1 require the notification, within two months, or as soon as reasonably possible, of certain national phase processing events to the International Bureau.

Prior to this amendment to the Regulations Offices had been requested to optionally provide National Phase Entry Information on a "best efforts" basis; this amendment represents a strengthening of the data collection process to respond to the needs to improve the data quality and coverage within the data that. The data collected is made available to the general public and the commercial patent information providers in a number of ways, including packaged data sets and internet-based queries for aggregate or individual application data.

This document describes the technical format for transmission of PCT national phase entry data using PCT-EDI Technical Services as the required means of transmittal.

RULE 95 INFORMATION AND TRANSLATIONS FROM DESIGNATED AND ELECTED OFFICES

The following updates to the PCT regulations, the revised PCT Rule 95.1 enters into force on July 1, 2017:

"95.1 INFORMATION CONCERNING EVENTS AT THE DESIGNATED AND ELECTED OFFICES

Any designated or elected Office shall notify the International Bureau of the following information concerning an international application within two months, or as soon as reasonably possible thereafter, of the occurrence of any of the following events:

(i) following the performance by the applicant of the acts referred to in Article 22 or Article 39, the date of performance of those acts and any national application number which has been assigned to the international application;

(ii) where the designated or elected Office explicitly publishes the international application under its national law or practice, the number and date of that national publication;

(iii) where a patent is granted, the date of grant of the patent and, where the designated or elected Office explicitly publishes the international application in the form in which it is granted under its national law, the number and date of that national publication."

NATIONAL PHASE ENTRY INFORMATION - DATA REQUIREMENTS

The data required for each event for each international application is described in the table below:

Data Item	Description
Office Code	WIPO Standard ST.3 code of the Office.
PCT IA Number	PCT International Application Number in the format PCT/XXYYYY/NNNNNN where XX is the receiving office code, YYYY is the year of PCT filing and NNNNNN is the application number assigned by the receiving Office.
National Application Number	Number attributed to the application at the time of national phase entry
Event type	"E" – National Phase Entry "P" – National Publication "G" – Grant "R" – Refusal "W" – Withdrawal "C" – Continuation of processing after a refusal or withdrawal "D" – Divisional application
Event date	YYYYMMDD: YYYY – year, MM – month (01-12), DD- day (01-31)
Publication associated with event	The national publication number associated with event categories P or (where relevant in the national system) G, without any WIPO ST.3 country code prefix.
Kind code of publication associated with event	The kind code of the publication number above, typically A or A1 for category P and B or B1 for category G.

Offices are required to send the first five data items specified for each event for PCT applications entering, or being processed in, the national phase:

- Office Code,
- International application number,
- National application number,
- Event type
- Event date.

In addition to these elements which are required in every case, Offices are required to send the relevant national publication numbers and publication kind codes when applicable to the relevant event.

PCT Rule 95.1 requires the information to be sent for event types E, P and G. Offices are encouraged also to send details of event types R and W where this is practical.

In some jurisdictions, there may be several different events which could be associated with the relevant code or events of a similar nature which are described in different terms. This standard aims to collect only the most basic information required to identify whether an international application is subject to national phase processing or has been granted and to allow the relevant national publications to be identified. As such, the following guidelines are offered as to how to map national phase events to the different event codes:

Category E: This should preferably be the date on which the actions required by Article 22 or 39 have been completed, even if this is prior to the expiration of the relevant time limit and the applicant has not requested national processing to begin early.

Category P: This should represent the main national publication. Republications need not be notified unless to correct a major error, such that the original publication was not effective.

Category G: In some Offices, grant involves several stages with distinct dates. For Offices which formally publish the patent as granted, it is recommended that this date report the date of the publication of the granted patent. Where no publication occurs, it is recommended that the date reflect the most significant date notified to the public, such as publication of the notice of grant in the national gazette.

Category R: This is intended to include any process where the Office takes an active decision not to grant a patent.

Category W: This is intended to include both active withdrawal by the applicant and any case where the application has lapsed, for example due to failure to respond to an examiner's report or failure to pay a fee.

Category C: This is intended to represent continuation of processing after a category R or W code has been sent, for example, because a decision to refuse has been successfully appealed or else fees have been paid to restore an application which had lapsed due to failure to meet a time limit.

Category D: This is intended to represent the division the initial national entry into a second, or subsequent, application with a new national application number after a category E code has been sent, for example, because an international application contained more than one invention, in the national phase it is divided into one application for each of the inventions described. Following such an entry, further events may be recorded for both the original and subsequently file national applications.

Note: the Category N event code has been retired as the IB intends to calculate in an consistent manner the non-entry of international applications; the Category X has also been retired.

EVENT TYPE BASED DATA UNITS

Where an Office is not in a position to provide data for all event types listed above, the Office must provide the data for the subset of event types for which it has data available.

In the interests of data integrity Offices are required to send data in files named according to sequence number, and include in each file, all events for the set of applications for which new events have occurred since the previous transmission; the International Bureau will replace all events in its database for the set of international applications in a given file with those included in the file.

DATA FORMATS

Offices are required to send data using either the Comma Separated Values (CSV) format or the Extensible Markup Language (XML) format.

NATIONAL PHASE ENTRY INFORMATION - CSV FILE SPECIFICATION

The standard CSV format specification:

CSV is a file format used as a portable representation of a database. Each line is one entry or record and the fields in a record are separated by commas (some rare cases use semicolons). Commas may be preceded or followed by arbitrary space and/or tab characters which are ignored.
If a field includes a comma or a new line, the whole field must be surrounded with double quotes. When the field is in quotes, any quote literal must be escaped by \". Backslash literals must be escaped by \\. Otherwise a backslash and the character following it will be treated as the following character, ie."\n" is equivalent to newline. Other escape sequences are "\n\r\t\f". Text that comes after quotes that have been closed but come before the next comma will be ignored.
Empty fields are returned as a String of length zero: "". The following line has four empty fields and two non-empty fields in it. There is an empty field on each end, and two in the middle.

```
,second,, ,fifth,
```

Blank lines are always ignored. Other lines will be ignored if they start with "#" or "!" as these characters denote comments.

PCT national phase entry CSV file format

Offices must provide one record for each event for each international application entering the National Phase.

Each record has 7 fields, all of which must be present in each record (noting that while positions 6 and 7 may be empty they must always be present).

Field Position	Field Name	Field Specification	Example value
1	Office code	The ST.3 uppercase Office code for the National Office sending the data	KR
2	PCT international application number	The IA number applicable to the record supplied in ST.10/C format (full "PCT/" preamble followed by office code and four digit year and then a "/" and the six digit number. i.e. "PCT/ROYYYY/NNNNNN")	PCT/JP2001/123456

3	National application number	The fully formatted National Application number. If this field contains commas it should be enclosed in quotes.	10200370000001
4	Event type	This field has seven possible values in uppercase only: "E" – National Phase Entry "P" – National Publication "G" – Grant "R" – Refusal "W" - Withdrawal "C" – Continuation of processing after a refusal or withdrawal "D" – Divisional application	E
5	Event date	The date associated with the event type in the format: "yyyymmdd"	20040124
6	Publication number	The national publication number associated with event categories P or (where relevant in the national system) G, without any WIPO ST.3 country code prefix. The fully formatted national publication number without the WIPO ST.3 country code prefix. If this field contains commas it should be enclosed in quotes. Where no national publication was effected on the date mentioned for a particular record an empty field must be included in the record.	10/2004/123456
7	Kind code of publication	The kind code of the publication number above, typically A or A1 for category P and B or B1 for category G. Where no national publication number was effected on the date mentioned for a particular record an empty field must be included in the record. (possible values are two character alpha-numeric codes [A..Z|1..0] [A..Z|1..0])	A1

An example CSV file is shown in Appendix I.

THE PCT NATIONAL PHASE ENTRY XML FILE FORMAT

XML Files are specified using document type definitions (DTDs) or XML Scheme Definitions (XSDs). The ST.36 compliant XML DTD is attached in Appendix II and sample XML data is attached in Appendix III.

DATA TRANSMISSION METHODS

For transmitting the data to the IB, transmission over the Internet using WIPO's PCT-EDI service is required; more information is given in Appendix IV.

NATIONAL PHASE ENTRY DATA TRANSMISSION FREQUENCY

Data updates must be transmitted to WIPO on a daily, weekly or monthly basis. Transmission on a monthly basis is the least preferred of these options. Whichever the transmission method the strict sequence of file numbers must be maintained to ensure that they are processed in the correct order.

NATIONAL PHASE ENTRY DATA BACK-FILE PROVISION

Where practical, Offices are requested to provide an initial national phase entry data extract with back-file data from 1998 onwards. The back-file data may need to be broken into units of 6 months or one year, depending on data volumes.

APPENDIX I

PCT NATIONAL PHASE ENTRY DATA- CSV FILE FORMAT EXAMPLES

```
KR,PCT/EP2001/007412,1020037000001,E,20010628,,
KR,PCT/EP2001/007412,1020037000001,P,20030102,1020047000010,A1
KR,PCT/EP2001/007412,1020037000001,W,20030111,,
KR,PCT/JP2001/005769,1020037000002,E,20030102,,
KR,PCT/JP2001/005769,1020037000002,P,20030206, 1020047000017,A1
KR,PCT/DE2002/001631,1020037000003,E,20020507,,
KR,PCT/DE2002/001632,1020037000007,E,20020507,,
KR,PCT/EP2001/007125,1020037000004,E,20010622,,
KR,PCT/EP2001/007125,1020037000004,P,20030303,1020047000099,A1
KR,PCT/EP2001/007125,1020037000004,G,20040621,,
KR,PCT/JP2002/004378,1020037000006,E,20020501,,
KR,PCT/JP2002/004378,1020037000006,P,20030102, 1020047000107,A1
KR,PCT/JP2002/004378,1020037000006,R,20030202,,
KR,PCT/JP2002/004378,1020037000006,C,20030301,,
KR,PCT/JP2002/004378,1020037000998,D,20030202,,
KR,PCT/JP2002/004378,1020037000999,D,20030202,,
KR,PCT/JP2002/004378,1020037000998,P,20030202,102047000201,A1
KR,PCT/JP2002/004378,1020037000998,G,20040110, 102047000201,B
KR,PCT/JP2002/004378,1020037000999,P,20030209,102047000202,A1
KR,PCT/JP2002/004378,1020037000999,G,20040124, 102047000202,B
```

In the example above:
- Application PCT/EP2001/007412 has entered the national phase, was published and then withdrawn
- application PCT/JP2001/005769 has entered National Phase and been published only;
- application PCT/DE2002/001631 has entered National Phase only;
- application PCT/DE2002/001632 has entered National Phase only;
- application PCT/EP2001/007125 has entered the national phase has been published and has been granted and,
- application PCT/JP2002/004378 has entered National Phase, was refused, continued following an appeal, divided with two further application numbers, both of which were published and then granted.

It should be noted that the data above is not real application data.

APPENDIX II

A. XML DTD for PCT National Phase Entry data

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

<!--Generated by Turbo XML 2.4.1.100.-->

<!--
************************************************************************************
* national-phase-information    2016  DECEMBER                                    *
************************************************************************************
*  PUBLIC "-//WIPO//NATIONAL PHASE INFORMATION1.0//EN" "wo-national-phase-
information-v1-5.dtd"  *
************************************************************************************
*  http://www.wipo.int/pct/edi/ed/dtd/wo-national-phase-information-v1-5.dtd      *
************************************************************************************
* contacts:                                                      *
*   WIPO: Peter Waring                                              *
*   WIPO: Asahi Hasebe                                             *
************************************************************************************
* revision history                                                               *
************************************************************************************
* April 2005                                                                       *
..released version number as v1.4                                                  *
* December 2016                                                                    *
..released version number as v1.5                                                  *
.. v1.5 is NOT backwardly compatible with v1.4                                     *
************************************************************************************


 ************************************************************************************
 *  DTD FOR NATIONAL PHASE DATA DOCUMENT                                          *
 *  Transfer Wrapper containing                              *
 *  National App No with IA No and then National event data.                  *
 *  This is DTD constructed in compliance with ST.36.                     *
 *                                                *
 * ROOT ELEMENT wo-national-phase-information                        *
 *                                            *
 *  All elements are taken from ST.36 ICE (International Common Elements) or       *
 *  wo-specific elements defined in accordance with ST.36.                    *
 *                                            *
 *  DTD redesigned for simple processing / obvious match with CSV and table  *
 *                                        *
 ************************************************************************************
 *                                            *
 ************************************************************************************
 * SECTION I - DECLARATION, ENT. REFs, ISO sets, etc                        *
 ************************************************************************************
-->
<!ELEMENT wo-national-phase-information (wo-national-office-event*)>
<!ATTLIST wo-national-phase-information
                        country        CDATA              #REQUIRED
                        file-name      CDATA                 #REQUIRED
                        date-produced  CDATA                 #REQUIRED
>
```

```
<!--
The following attribute usage is expected
      country – this attribute is required to indicate the country of entry events using the ST.3
code
      file-name – this attribute is required and must match the actual name of the file including
the sequence number.
      date-produced – this attribute is required to indicate the processing date of the file creation

-->


<!--
  National Office event
  Enhanced entry data requires additional information on the national phase applications.
  Status code is made up of as follows;
   "E" – National Phase Entry
   "P" – National Publication
   "G" – Grant
   "R" – Refusal
   "W" - Withdrawal
   "C" – Continuation of processing after a refusal or withdrawal
   "D" – Divisional application
-->
<!ELEMENT wo-national-office-event EMPTY >
<!ATTLIST wo-national-office-event  id        ID            #IMPLIED
  office-code CDATA #REQUIRED
  international-application CDATA #REQUIRED
  national-application CDATA #REQUIRED
  event-type  ( E | P | R | G | W | C | D )  #REQUIRED
  event-date CDATA #REQUIRED
  publication-number CDATA #IMPLIED
  publication-kind-code CDATA #IMPLIED >
```

APPENDIX III

XML EXAMPLE INCLUDING EVENT TYPES

This example reflects the same sample data used in the CSV example in main body of the document.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE wo-national-phase-information SYSTEM "wo-national-phase-information-v1-5.dtd">
<wo-national-phase-information country='KR' file-name= 'KR-IB-npsd-000001.zip' date-produced = '14122016' >
  <wo-national-office-event office-code='KR'  international-application = 'PCT/EP2001/007412' national-application = '1020037000001' event-type = 'E' event-date = '20010628'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/EP2001/007412' national-application = '1020037000001' event-type = 'P' event-date = '20030102'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/EP2001/007412' national-application = '1020037000001' event-type = 'E' event-date = '20030102' publication-number='1020047000010' publication-kind-code='A1' />
  <wo-national-office-event office-code='KR' international-application = 'PCT/EP2001/007412' national-application = '1020037000001' event-type = 'W' event-date = '20030111'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/JP2001/005769' national-application = '1020037000002'  event-type = 'E' event-date=' 20030102'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/JP2001/005769' national-application = '1020037000002' event-type= 'P'  event-date=' 20030206' publication-number='1020047000017' publication-kind-code='A1' />
  <wo-national-office-event office-code='KR' international-application = 'PCT/DE2002/001631' national-application = '1020037000003'  event-type = 'E' event-date=' 20020507'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/DE2002/001632' national-application = '1020037000007'  event-type = 'E' event-date=' 20020507'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/EP2001/007125' national-application = '1020037000004'  event-type = 'E' event-date=' 20010622'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/EP2001/007125' national-application = '1020037000004'  event-type = 'P' event-date=' 20030303' publication-number='1020047000099' publication-kind-code='A1' />
  <wo-national-office-event office-code='KR' international-application = 'PCT/EP2001/007125' national-application = '1020037000004'  event-type = 'G' event-date=' 20040621'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/JP2002/004378' national-application = '1020037000006'  event-type = 'E' event-date=' 20020501'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/JP2002/004378' national-application = '1020037000006'  event-type = 'E' event-date=' 20030102' publication-number='1020047000107' publication-kind-code='A1' />
  <wo-national-office-event office-code='KR' international-application = 'PCT/JP2002/004378' national-application = '1020037000006'  event-type = 'R' event-date=' 20030202'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/JP2002/004378' national-application = '1020037000006'  event-type = 'C' event-date=' 20030301'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/JP2002/004378' national-application = '1020037000998'  event-type = 'D' event-date = '20030302'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/JP2002/004378' national-application = '1020037000999'  event-type = 'D' event-date=' 20030302'/>
  <wo-national-office-event office-code='KR' international-application = 'PCT/JP2002/004378' national-application = '1020037000998'  event-type = 'P' event-date=' 20030402' publication-number='1020047000201' publication-kind-code='A1'/>
```

```
   <wo-national-office-event office-code='KR' international-application = 'PCT/JP2002/004378'
national-application = '1020037000999'  event-type = 'P' event-date=' 20040110' publication-
number='1020047000202' publication-kind-code='A1'/>
   <wo-national-office-event office-code='KR' international-application = 'PCT/JP2002/004378'
national-application = '1020037000998'  event-type = 'G' event-date=' 20030409' publication-
number='1020047000201' publication-kind-code='B'/>
   <wo-national-office-event office-code='KR' international-application = 'PCT/JP2002/004378'
national-application = '1020037000999'  event-type = 'G' event-date=' 20040124' publication-
number='1020047000202' publication-kind-code='B'/>
</wo-national-phase-information>
```

APPENDIX IV

PCT-EDI TRANSMISSION METHOD

The International Bureau offers an electronic Data Interchange server based on secure FTP. How to upload national phase entry information should be uploaded by transferring individually each national Phase entry file to the "Upload" folder using the Office PCT-EDI account is described in the one-page "How to transfer PCT National Phase Information via PCT EDI service 10.doc" document available on the PCT-EDI web page. Additional information can be found in the documents on the following webpage:

http://www.wipo.int/pct/edi/en/index.html

DATA WRAPPER NAMING CONVENTION

A "wrapper file" in standard ZIP format must be created for each national phase entry data file transmitted to the International Bureau. The name of a wrapper file is composed of five consecutive parts separated by dashes (based on the document wrapper naming convention of Minimal Specifications for Electronic PCT Document Exchange):

1)   The upper case WIPO Standard ST.3 code of the transmitting office
2)   The upper case WIPO Standard ST.3 code of the International Bureau (IB)
4)   The document type code, npsd, in lower case: npsd representing national phase entry data
5)   A numeric string NNNNNN to give the file its processing sequence number (files will be processed sequentially in order with no gaps): NNNNNN is a number right justified and padded with leading zeroes

Example of a correctly named ZIP file for transfer: KR-IB-npsd-000001.zip