PCT IS Division

Data Format Specifications
for the
Collection of PCT National Phase Information

Version Number 3.0
May 14, 2007

WORLD INTELLECTUAL PROPERTY ORGANIZATION

GENEVA

Document Information

Document title:	Data Format Specifications for the Collection of PCT National Phase Information
Document file name:	edi_nat_phase_entry_format_reqs_3.0.doc
Issued by:	Messrs. Kader Taibi and Young-Woo Yun
Reviewed by	Messrs. James Fullton, William Meredith and Peter Waring
Issue Date:	May 14, 2007
Status:	Approved

References

1. PCT EDI Cooperation Project.

2. Minimal Specification for Transmitting Documents to the IB.

3. ST 36 WIPO Standards.

4. How to transfer PCT National Phase Information via PCT EDI service 10.doc.

Table of Contents

1. Introduction

The purpose of this document is to describe the technical requirements for provision of PCT national phase entry data by Intellectual Property (IP) Offices to the International Bureau (IB) of WIPO.

For national phase entry information there is an increasing demand for IP Information from offices, analysts, policy makers and others. To respond to this demand, WIPO publishes in the "PatentScope" web site, under the PCT Rule 86.1(iv), the information collected from designated/elected Offices concerning the PCT national/regional phase entry. It is intended that the data would be 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.

In parallel, WIPO has initiated the PCT Electronic Data Exchange Project (PCT-EDI) in order to facilitate the internet-based exchange of a wide range of electronic documents and data between WIPO and IP Offices.

This document describes the technical format for exchange of PCT national phase entry data using PCT-EDI Technical Services as the preferred means of transmittal; email is an equally acceptable means of transmission.

2. Data Required

National Phase Entry Information

National Phase Entry Information - Data Requirements
The data required is described in the table below:

Data Item	Description
Office Code	WIPO standard ST.3 code of the reporting 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 sequential application number assigned by the receiving office.
National Application Number	Number attributed to the application at the time of national phase filing (not the publication number)
Date of National Phase Entry	
Date of Publication	
Date of Grant	
Date of Refusal	
Date of Withdrawal	

All participating offices are requested to send the first four data items for all PCT national phase entries: Office Code, PCT IA Number, National Application Number and Date of National Phase Entry.

Offices are requested to send the dates of events after national phase entry (publication, grant, refusal and/or withdrawal) for those PCT applications for which one, or more, of the events has occurred since the last update.

Example 1. Minimum data:

Office Code	PCT IA Number	National application number	Date of National Phase Entry	Publication Date	Grant Date	Refusal Date	Withdrawal Date
KR	PCT/JP2001/000731	1020037524420	20030103				
KR	PCT/EP2001/007201	1020037523421	20030103				
KR	PCT/JP2001/000731	1020037524420	20030103				
KR	PCT/EP2001/007201	1020037523421	20030103				
KR	PCT/US2001/017384	1020037523403	20030103				
KR	PCT/CN2001/007384	1020037523493	20040103				

Example 2. Complete data:

Office Code	PCT IA Number	National application number	Date of National Phase Entry	Publication Date	Grant Date	Refusal Date	Withdrawal Date
KR	PCT/JP2001/000731	1020037524420	20030103	20030601			
KR	PCT/EP2001/007201	1020037523421	20030103	20030601			
KR	PCT/JP2001/000731	1020037524420	20030103	20030614	20040204		
KR	PCT/EP2001/007201	1020037523421	20030103	20030614			20030301
KR	PCT/US2001/017384	1020037523403	20030103				
KR	PCT/CN2001/007384	1020037523493	20030103	20030601		20031203	

Note that when the complete data is provided, a transmission may contain new event dates (publication, grant, refusal or withdrawal) for applications already provided in earlier transmissions.

3. Data Formats

National Phase Entry Information

Offices are requested to send data in one of the following two formats:

Either
1. Comma Separated Values (CSV). CSV format data files can be easily generated by commonly available software such as Microsoft Excel and OpenOffice.org Calc,

Or
2. Extensible Markup Language (XML). The ST.36 compliant XML DTD and sample XML data are attached in Appendix II.

In the case where it is not possible to furnish the data to WIPO in one of the two formats described in this document, an office should consult with WIPO to determine whether a bi-lateral agreement can be used to define an alternative 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.

The PCT national phase entry CSV file format is as follows:

Each record has 5 fields, all of which must be present on each line.

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 IA 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 five possible single character values in uppercase only: E – National Phase Entry (National Filing) P – National publication R – National Refusal G - National Grant W – National withdrawal.	E
5	Date of Action	The date associated with the event type in the format: "yyyymmdd"	20040124

It should be noted that multiple rows for an application are permitted. i.e. in the case of an application that has entered national phase and has been published (but no grant, refusal or withdrawal has been processed) two rows are expected, one with status code E and one with status code P. Additionally it is expected that for each application that is sent there should always be a row with a Status Code of E.

An example CSV file is shown in Appendix III.

National Phase Entry Data back-file provision

Where practical, participating 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.

National Phase Entry Data update frequency

The updated data shall be transmitted to WIPO on a monthly or quarterly basis. Monthly transmission is preferred.

4. Data Transmission Methods

Two methods are supported for transmitting the data to the IB.

Either
1. Transmission over the Internet using WIPO's PCT-EDI service; this is the preferred means of transmittal,

Or
2. Email attachment sent to ipstats.mail@wipo.int.

The use of the simple, flexible WIPO PCT EDI service is recommended. More information is given in Appendix I.

Appendix I

A. PCT-EDI Transmission Method

The International Bureau offers an electronic Data Interchange server based on secure FTP. How to upload national phase entry information 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 website:

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

B. Data Wrapper Naming Convention

A "wrapper file" in standard ZIP format must be created for each set of 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 start date followed by the end date applicable to the national phase entry data following the "YYYYMMDD-YYYYMMDD" format in local office time
3) The document type code, npsd, in lower case: npsd representing national phase entry data
4) A numeric string NNNNNN to make the filename unique within its directory: NNNNNN is a number right justified and padded with leading zeroes
5) The upper case ISO639 code representing the language of the document or XX if unknown.

Example of a correctly named ZIP file for transfer: *KR-20050201-20050228-npsd-000001-EN.zip*

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    2005  APRIL                                                  *
**********************************************************************************************
*  PUBLIC "-//WIPO//NATIONAL PHASE INFORMATION1.0//EN" "wo-national-phase-information-v1-4.dtd"  *
**********************************************************************************************
* http://www.wipo.int/pct/edi/en/dtd/wo-national-phase-information-v1-4.dtd         *
**********************************************************************************************
* contacts:                                                                                  *
*   WIPO: Peter Waring                                                                        *
*   WIPO: Young-Woo Yun                                                                       *
**********************************************************************************************
* revision history                                                                           *
**********************************************************************************************
* April 2005                                                                                 *
..released version number as 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 and uses elements                    *
   *   specified in the ICE. It is intended for use as an inter office communication         *
   *   where documents have been reclassified in IPC version 8.                              *
   *                                                                                          *
   *   ROOT ELEMENT wo-nataional-phase-information                                           *
   *                                                                                          *
   *   All elements are taken from ST.36 ICE (International Common Elements) or               *
   *   wo-specific elements defined in accordance with ST.36.                                *
   **********************************************************************************************
   *                                                                                          *
   **********************************************************************************************
   * SECTION I - DECLARATION, ENT. REFs, ISO sets, etc                                        *
   **********************************************************************************************
-->
<!ELEMENT wo-national-phase-information (wo-national-pct-reference+)>
<!ATTLIST wo-national-phase-information id               ID                        #IMPLIED
                                        country          CDATA                     #REQUIRED
                                        lang             CDATA                     #REQUIRED
                                        information-type (complete | incremental ) #REQUIRED
                                        period-start-date CDATA                    #REQUIRED
                                        period-end-date  CDATA                     #REQUIRED
                                        dtd-version      CDATA                     #IMPLIED
                                        file             CDATA                     #IMPLIED
                                        file-reference-id CDATA                    #IMPLIED
                                        date-produced    CDATA                     #IMPLIED >
<!--The period-start-date and period-end-date attributes are used for a certain period of time
covered by
    this entry data.  For example, if the entry data includes information on national applications
    filed from 2004.01.01 to 2004.12.31, the period-start-date value will be 20040101 and
    the period-end-date value will be 20041231.

The information-type attribute is used to show if this is an incremental file within a series of
files
    or a transfer of all available information at the time of extraction.
-->


<!--    Filed national Application;
-->
<!--
Application reference information: application number, country
(INID 21, ST.32:B210)

international-Application-number format example PCT/KR2002/123456
```

```
-->

<!ELEMENT wo-national-pct-reference (wo-ia-number, application-reference, wo-national-filing-date,
wo-national-office-event*)>
<!ATTLIST wo-national-pct-reference  id                                    ID            #IMPLIED  >

<!ELEMENT wo-ia-number (application-reference)>

<!--     Filed national Application;
-->
<!ELEMENT application-reference (document-id)>
<!ATTLIST application-reference  id       ID     #IMPLIED
                                 appl-type CDATA  #REQUIRED >

<!ELEMENT wo-national-filing-date (date)>

<!--
  Enhanced entry data requires additional information on the national phase applications.
  Status code is made up of as follows;
     P: Published
    G: Granted
    R: Refused
    W: Withdrawn
-->
<!ELEMENT wo-national-office-event (wo-event-date)>
<!ATTLIST wo-national-office-event  id           ID              #IMPLIED
                                    event-type  (P | R | G | W )  #REQUIRED>

<!--
  The wo-event-date is the date when the wo-national-office-event happened to the application.
-->
<!ELEMENT wo-event-date (date)>

<!ELEMENT document-id (country, doc-number, kind?, name?, date?)>
<!ATTLIST document-id  lang CDATA  #IMPLIED >
<!--
  Country: use ST.3 country code; e.g. KR, JP, US, etc.
  Also includes EP, WO
-->
<!ELEMENT country (#PCDATA)>

<!--
  The number of the referenced patent (or application) document
-->
<!ELEMENT doc-number (#PCDATA)>

<!--
  Document kind code; e.g., A1
    (INID 13, ST.32:B130)
-->
<!ELEMENT kind (#PCDATA)>

<!--
  Name:
  If no distinction or detail can be given.
  Also to be used for: personal (natural person) and corporate (legal entity) names
-->
<!ELEMENT name (#PCDATA)>

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

<!--
  Date: components of a date.  Format is YYYYMMDD.
-->
<!ELEMENT date (#PCDATA)>
```

B. *Sample XML data for PCT National Phase Entry data*

Including *Complete* Data for example 2

```
<?xml version = "1.0" encoding = "UTF-8"?>
<!DOCTYPE wo-national-phase-information SYSTEM "wo-national-phase-information-v1 4.dtd">
<wo-national-phase-information country = "KR" lang = "en" information-type = "incremental" period-start-date = "20030101"
period-end-date = "20031231">
        <wo-national-pct-reference>
                <wo-ia-number>
                        <application-reference appl-type = "">
                                <document-id>
                                        <country/>
                                        <doc-number>PCT/JP2001/00731</doc-number>
                                </document-id>
                        </application-reference>
                </wo-ia-number>
                <application-reference appl-type = "">
                        <document-id>
                                <country>KR</country>
                                <doc-number>1020037524420</doc-number>
                        </document-id>
                </application-reference>
                <wo-national-filing-date>
                        <date>20030103</date>
                </wo-national-filing-date>
                <wo-national-office-event event-type = "P">
                        <wo-event-date>
                                <date>20030601</date>
                        </wo-event-date>
                </wo-national-office-event>
                <wo-national-office-event event-type = "W">
                        <wo-event-date>
                                <date>20031004</date>
                        </wo-event-date>

                </wo-national-office-event>
        </wo-national-pct-reference>
        <wo-national-pct-reference>
                <wo-ia-number>
                        <application-reference appl-type = "">
                                <document-id>
                                        <country/>
                                        <doc-number>PCT/EP2001/007201</doc-number>
                                </document-id>
                        </application-reference>
                </wo-ia-number>
                <application-reference appl-type = "">
                        <document-id>
                                <country>KR</country>
                                <doc-number>1020037523421</doc-number>
                        </document-id>
                </application-reference>
                <wo-national-filing-date>
                        <date>20030103</date>
                </wo-national-filing-date>
                <wo-national-office-event event-type = "P">
                        <wo-event-date>
                                <date>20030601</date>
                        </wo-event-date>
                </wo-national-office-event>
                <wo-national-office-event event-type = "G">
                        <wo-event-date>
                                <date>20040204</date>
                        </wo-event-date>
                </wo-national-office-event>
        </wo-national-pct-reference>
</wo-national-phase-information>
```

Appendix III

A. PCT National Phase Entry data- CSV file format example

```
KR,PCT/EP2001/007412,1020037000001,E,20010628
KR,PCT/EP2001/007412,1020037000001,P,20030102
KR,PCT/EP2001/007412,1020037000001,W,20030104
KR,PCT/JP2001/005769,1020037000002,E,20030102
KR,PCT/JP2001/005769,1020037000002,P,20030206
KR,PCT/DE2002/001631,1020037000003,E,20020507
KR,PCT/EP2001/007125,1020037000004,E,20010622
KR,PCT/EP2001/007125,1020037000004,P,20030303
KR,PCT/EP2001/007125,1020037000004,G,20040621
KR,PCT/JP2002/004378,1020037000006,E,20020501
KR,PCT/JP2002/004378,1020037000006,P,20030102
KR,PCT/JP2002/004378,1020037000006,R,20030202
```

In the example above application PCT/DE2002/001631 has entered National Phase only, application PCT/JP2001/005769 has entered National Phase and been published etc. It should be noted that the data above is not real application data.