



**Weekly Published PCT Data (Bibliographic Data) through FTP site - <ftp://wwwftp.wipo.int>  
Content and File Structure (as of July 1st, 2008)**

Content	Directory or File Name	Content Description
Publication Indexation List	PCT_{\$gazno}.lst	Listing of all IAs in the Publication
Bibliographic Data in TIFF and XML Formats + Drawings and Abstracts in TIFF Format	PCT_{\$gazno}.tar.gz	When extracted the archive file contains one directory named <b>{\$gazno}</b> which contains: <ul style="list-style-type: none"> <li>• index.lst (publication list)</li> <li>• {\$gazno}-gazette-section-ii.xml (See Annex B)</li> <li>• One directory for each IA in the publication <b>may contain</b>: <ul style="list-style-type: none"> <li>– {\$Iano}_{\$pubdate}/wo-published-application.xml</li> <li>– {\$Iano}_{\$pubdate}/000001.tif ... nnnnnn.tif (front pages)</li> <li>– {\$Iano}_{\$pubdate}/Pag.lst (legacy indexation)</li> <li>– {\$Iano}_{\$pubdate}/{\$Iano}_{\$pubdate}/_pf_fp.g4</li> <li>– {\$Iano}_{\$pubdate}/{\$Iano}_{\$pubdate}/_gz_en.g4</li> <li>– {\$Iano}_{\$pubdate}/{\$Iano}_{\$pubdate}/_gz_fr.g4</li> <li>– {\$Iano}_{\$pubdate}/{\$Iano}_{\$pubdate}/_pf_ab.g4</li> </ul> </li> </ul>
Bibliographic Data in XML Format	PCT_{\$gazno}_xml.lst PCT_{\$gazno}_xml.tar.gz	When extracted the archive file contains one directory named <b>{\$gazno}</b> which contains: <ul style="list-style-type: none"> <li>• index.lst</li> <li>• One directory for each IA in the publication containing: <ul style="list-style-type: none"> <li>– {\$Iano}_{\$pubdate}/wo-published-application.xml</li> </ul> </li> </ul>
Index List	Index.lst	Index of International Applications in the Weekly Publication



**Weekly Published PCT Data (Bibliographic Data + Data Images) through FTP site - <ftp://wwwftp.wipo.int/ftpctest>  
Content and File Structure (as of July 1st, 2008)**

**NOTE: Access to the FTP server also provides you access to test data. This means access to complete published PCT international patent applications (bibliographic data, descriptions and claims) in image format and OCR full-text format. This is part of a pilot project to deliver complete PCT publication data by FTP. We are currently investigating several options to increase our bandwidth and capacity to support downloads by FTP because of the large volumes of data involved. As a result, this means of delivery of PCT data remains in a test phase.**

Content	Directory or File Name	Content Description
PCT International Applications OCR in XML Format	ftpctest/\${gazno}.zip (e.g. 2007-30.zip)	XML Documents containing the OCR results for description and claims, if present, for each International Application (IA) in the weekly publication named: <b>/\${gazno}/doc/wo\${repubtype}/\${kindcode}/WOccynnnnnn\${repubtype}/\${kindcode}.zip</b> (NOTE: \$repubtype = "nw" for 1 <sup>st</sup> publications)  <b>index.xml</b> – index file of all ocr xml zip files in the publication identified by the document WO number.  Volumeid – batch identification (e.g. publication week)



**Weekly Published PCT Data (Bibliographic Data + Data Images) through FTP site - <ftp://wwwftp.wipo.int/ftptest>  
Content and File Structure (as of July 1st, 2008)**

Content	Directory or File Name	Content Description
PCT International Applications (Bibliographic Data plus Data Images)	The entire weekly publication is split into a number of archive files named: ftptest/\${gazno}/IBnnnn_\${gazno}_DONE_a.tar ftptest/\${gazno}/IBnnnn_\${gazno}_DONE_b.tar . . . ftptest/\${gazno}/IBnnnn_\${gazno}_DONE_z.tar	When extracted the archive files contain one or more zip files per International Application including: <ul style="list-style-type: none"> <li>• the Publication WAD named:  <b>IBnnnn_\${gazno}_DONE/documents/\${IAno}/WO\${Wono}_\${pubdate}.zip</b></li> <li>- the Publication WAD includes the following:               <ul style="list-style-type: none"> <li>- data images 000001.tif ... nnnnnn.tif (one tif image per A4 page)</li> <li>- drawing images (where present)                   <ul style="list-style-type: none"> <li>    \${IAno}_\${pubdate}_pf_fp.g4 (publication language)</li> <li>    \${IAno}_\${pubdate}_gz_fr.g4 (French drawing translation)</li> <li>    \${IAno}_\${pubdate}_gz_en.g4 (English drawing translation)</li> </ul> </li> <li>- wo-published-application.xml (bibliographic data and document indexation)</li> <li>- wo-published-application-ocr.pdf (searchable OCR results of the description and claims where present)</li> <li>- packagedata-pkda.xml (Content description of the Publication WAD)</li> <li>- Pag.lst (legacy indexation file)</li> </ul> </li> <li>• and 0 or more optional files (see Annex A) named: {IAno}_\${doctype}_\${docid}.zip</li> </ul> The final archive file contains, in addition to the IA files, a list of all IAs published in a week, and the Notices related to published IAs in XML format  <b>Ctrl.lst</b> (Index of all International Applications in the publication) <b>\${gazno}-gazette-section-ii.xml</b> (see Annex B)
Sequence Listing Index	ftptest/SL.lst	Index of published IAs, total page count and sequence listing (SL) page range
Index List	ftptest/Index.lst	Index of International Applications in the Weekly Publication



Weekly Published PCT Data (Bibliographic Data + Data Images) through FTP site - <ftp://wwwftp.wipo.int/ftptest>  
Content and File Structure (as of July 1st, 2008)

### Annex A – Optional Document Types

Each IA directory (i.e. IBnnnn\_{\$gazno}\_DONE/documents/{\$IAno}) may contain, in addition to the Publication WAD, zero or more optional documents. The document WADs are named: {\$IAno}\_{\$doctype}\_{\$docid}.zip

Document Description	doctype
Biological Material	biod
Request for Publication of Refused Request for Rectification	aprr
Publication of Late Submitted Request to Correct/Add Priority Claims	rlpc
Declaration	decl
Request for Publication of Information regarding Defects in Priority Claims	rppc
Statement of Reasons (Restoration of Priority Claims)	stpr
Form PCT/RO/114 – Notification on Decision of Confirmation of Incorporation by Reference of Element or Part	ro114
Form PCT/IB/318 – Notification Relating to Priority Claim	Ib318



**Weekly Published PCT Data through FTP site  
Content and File Structure (as of July 1st, 2008)**

**Annex B – Notices related to Published International Applications**

Content	Directory or File Name	Content Description
<p>Notices related to published International Applications (IAs):</p> <ol style="list-style-type: none"> <li>1. Announcements of the:               <ol style="list-style-type: none"> <li>a) Withdrawal of IAs after international publication</li> <li>b) Declaration by the Receiving Office that IAs are considered withdrawn</li> <li>c) Withdrawal of designations after international publication</li> <li>d) Withdrawal of priority claims before completion of the technical preparations</li> <li>e) Withdrawal of priority claims after completion of the technical preparations</li> </ol> </li> <li>2. Demands for International Preliminary Examination:               <ol style="list-style-type: none"> <li>a) (made in respect of IAs filed before January 1, 2004)</li> <li>b) (made in respect of IAs filed on or after January 1, 2004)</li> </ol> </li> <li>3. Corrigendum</li> </ol>	<p style="text-align: center;">\${gazno}-gazette-section-ii.xml</p>	<p>Announcements in compliance with the gazette-section-ii-v1.dtd</p>