OECD DATABASE ON TRIADIC PATENT FAMILIES

February 2015 edition

BACKGROUND INFORMATION

A patent family is defined as a set of patents taken in various countries for protecting a same invention, i.e. related patents are regrouped into a single record to derive a unique patent family. The OECD "Triadic" Patent Families consists in a set of patents filed for at the European Patent Office (EPO), the Japan Patent Office (JPO) and the United States Patent and Trademark Office (USPTO) that share one or more priority applications.

Note that the coverage of the Triadic Patent Families has been extended and now includes **patent applications published by the USPTO** from the early 2000s, and do not only refer to USPTO patent grants as in previous editions of the dataset.

The *EPO's Worldwide Statistical Patent Database* (PATSTAT, Autumn 2014) is the primary data source: PATSTAT database provides a harmonised and comparable set of information on EPO, JPO and USPTO patents.

DATABASE STRUCTURE & COVERAGE

The OECD Triadic Patent Families database (TPF) mostly rely on the INPADOC family table provided in PATSTAT, that is further consolidated in order to integrate patent families linkages using PCT information when existing. Data is presented according to a unique family identifier (Family_id) that is automatically generated for each new version of the database (it is therefore not possible to compare identifiers from one edition to another). The TPF data can be linked to PATSTAT recent editions, using the surrogate key Appln_id of each family member.

The OECD TPF dataset is provided as 9 distinct files, according to the levels of information on each family. Bibliographic information is provided for all triadic patent families, in sets of flat files based on UTF-8 format, using a pipe character ("|") field separator, text being enclosed with quotes "". The data is organised as a set of relational tables that can be connected using the unique Family identifier (Family_id). Mind that the Family_id is automatically generated for each version of the OECD TPF Dataset and should not be compared from one version to another. February 2015 edition now covers 1,243,221 distinct patent families, broken down by earliest priority date as shown in the table opposite.

REFERENCES

Martinez, C. (2010) "Insight into different types of patent families", STI Working Paper 2010/2, OECD, Paris.

OECD (2009) OECD Patent Statistics Manual, OECD, Paris

Dernis, H. & M. Khan (2004) "Triadic Patent Families Methodology", STI Working Paper 2004/2, OECD, Paris

RESTRICTIONS SOURCE & CONTACT

Please note that the raw data on OECD Triadic Patent Families are provided for research and analytical work. Make sure it is quoted as ***"OECD, Triadic Patent Families database, February 2015"*** when publishing the results of your analysis.

For further information about the OECD patent project, the methodology developed for building triadic patent families and access to pre-defined patent indicators *(families by technologies, by country, nowcasted figures...)*, please visit our web page at www.oecd.org/sti/ipr-statistics.

Comments and questions about this dataset should be sent to STI.Microdatalab@oecd.org.
For further information on EPO's PATSTAT, please contact patstat@epo.org.

Earliest priority date	First Family-id	Last Family-id	Number of Families *USPTO application*	Number of Families *USPTO grants*
All	1	1,243,221	**1,243,221**	983,481
2013	1,243,216	1,243,221	**6**	-
2012	1,233,737	1,243,215	**9,479**	487
2011	1,210,568	1,233,736	**23,169**	3,767
2010	1,162,171	1,210,567	**48,397**	12,421
2009	1,112,494	1,162,170	**49,677**	20,649
2008	1,062,393	1,112,493	**50,101**	27,564
2007	1,009,658	1,062,392	**52,735**	32,455
2006	952,888	1,009,657	**56,770**	36,163
2005	892,137	952,887	**60,751**	38,762
2004	831,319	892,136	**60,818**	40,158
2003	773,164	831,318	**58,155**	40,269
2002	717,145	773,163	**56,019**	41,042
2001	662,605	717,144	**54,540**	42,226
2000	606,966	662,604	**55,639**	45,757
1999	559,530	606,965	**47,436**	45,066
1998	514,599	559,529	**44,931**	43,651
1997	471,958	514,598	**42,641**	41,881
1996	432,165	471,957	**39,793**	39,386
1995	395,809	432,164	**36,356**	36,161
1994	362,746	395,808	**33,063**	32,990
1993	331,424	362,745	**31,322**	31,277
1992	300,984	331,423	**30,440**	30,419
1991	270,115	300,983	**30,869**	30,849
1990	237,563	270,114	**32,552**	32,539
1989	204,607	237,562	**32,956**	32,949
1988	174,224	204,606	**30,383**	30,376
1987	146,176	174,223	**28,048**	28,046
1986	121,791	146,175	**24,385**	24,383
1985	99,068	121,790	**22,723**	22,723
1984	78,042	99,067	**21,026**	21,026
1983	59,423	78,041	**18,619**	18,618
1982	43,457	59,422	**15,966**	15,966
1981	29,703	43,456	**13,754**	13,754
1980	17,832	29,702	**11,871**	11,871
1979	8,313	17,831	**9,519**	9,519
1978	1	8,312	**8,312**	8,311

DATABASE STRUCTURE

TPF_EPO
List of EPO patents in the family — *1,429,067 rows*

Family_id	Family identifier (from CORE table – repeated entry)
EPO_Nbr	EPO application number ("EPYYYYNNNNNNN")
Appln_id	Surrogate key - applications in PATSTAT, Autumn 2014
TItle	English title of the EPO filing

TPF_USPTO
List of USPTO patents in the family — *1,786,212 rows*

Family_id	Family identifier (from CORE table – repeated entry)
USPTO_Nbr	USPTO publication number ("US0NNNNNNN" for **USPTO Grants** "USYYYYNNNNNN" for **USPTO Applications**)
Appln_id	Surrogate key - applications in PATSTAT, Autumn 2014
TItle	English title of the USPTO patent

TPF_JPO
List of JPO patents in the family — *1,493,320 rows*

Family_id	Family identifier (from CORE table – repeated entry)
JPO_Nbr	JPO application number (not harmonised)
Appln_id	Surrogate key - applications in PATSTAT, Autumn 2014
TItle	English title of the JPO patent *only where an english translation is available*

TPF_PCT
List of PCT filings in the family — *744,227 rows*

Family_id	Family identifier (from CORE table – repeated entry)
PCT_Nbr	PCT publication number ("WOYYYYNNNNNN")
Appln_id	Surrogate key - applications in PATSTAT, Autumn 2014

TPF_PRIORITY
List of priority filings + other patent linkages — *2,972,879 rows*

Family_id	Family identifier (from CORE table – repeated entry)
Patent_nbr	Priority number or related patent application (Country code 'CC'+Application number)
Parent_Type	Identification of the patent linkages using PATSTAT tables: PRIO for Paris convention links (from APPLN_PRIOR table); CONTN for continuations in part or divisions (from APPLN_CONTN table) ; TECH for technical relations (from TECH_REL table) ; PCT for links based on PCT filings (using Internat_Appln_id from APPLN table)
Appln_id	Surrogate key - applications in PATSTAT, Autumn 2014

TPF_CORE
Summarised composition of families — *1,243,221 rows*

Family_id	Unique family identifier : 1 to 1,243,221 (from the earliest priority to the latest; mind that the identifier is automatically generated for each data release)
Count_prio	Total number of priority applications in the family (counts only the priority links Paris convention definition: count_prio=0 when other patent linkages are used – e.g. technical relations, divisions or continuations, PCT, etc.)
First_Prio	Earliest priority date – YYYYMMDD – when count_prio=0, the earliest date of patent family members,
Last_prio	Latest priority date – YYYYMMDD
USPTO_app_first	Earliest filing date at USPTO – YYYYMMDD
USPTO_app_Last	Latest filing date at USPTO – YYYYMMDD
USPTO_grant	Optional - Earliest date of grant at USPTO - YYYYMMDD
Count_USPTO	Total number of USPTO patents granted that belong to the family
EPO_app_first	Earliest filing date at EPO – YYYYMMDD
EPO_app_last	Latest filing date at EPO - YYYYMMDD
EPO_grant	Optional – Earliest date of grant at the EPO - YYYYMMDD
Count_EPO	Total number of EPO patent filings that belong to the family
JPO_app_first	Earliest filing date at JPO - YYYYMMDD
JPO_app_last	Latest filing date at JPO - YYYYMMDD
JPO_grant	Optional – Earliest date of grant at the JPO - YYYYMMDD
Count_JPO	Total number of JPO patent filings that belong to the family
PCT_app_first	Optional – Earliest PCT filing date - YYYYMMDD
Count_PCT	Optional – Total number of PCT filings that belong to the family, if any

TPF_IPC
list of IPC classes of each family member — *25,178,952 rows*

Family_id	Family identifier (from CORE table – repeated entry)
Appln_id	Surrogate key - applications in PATSTAT, Autumn 2014
Appln_auth	Patent application authority (EP/JP/US)
IPC	International Patent Classification, 8th edition
Count_IPC	Count of IPC classes per family member

TPF_INVENTORS
List of inventors[1] in the family - from USPTO patents — *3,822,010 rows*

Family_id	Family identifier (from CORE table – repeated entry)
Inventor	Name of the inventor
Address	City of residence of the inventor
Country	Country of residence of the inventor - 2 digits ISO code
Inventor_Count	Total number of inventors included in the family
Inventor_Share	Share[2] of each inventor in the family *(for fractional counts)*

TPF_APPLICANTS
List of applicants[1] in the family - from USPTO patents — *2,778,182 rows*

Family_id	Family identifier (from CORE table – repeated entry)
Applicant	Name of the applicant
Address	City of residence of the applicant
Country	Country of residence of the applicant - 2 digits ISO code
Applicant_Count	Total number of applicants included in the family
Applicant_Share	Share[2] of each applicant in the family *(for fractional counts)*

1. When missing, inventors (respectively applicants) were added to the applicant's table (resp. inventor's table).
2. Shares of applicants/inventors are provided to facilitate fractional patent counts: e.g. for a patent family co-invented by 1 French, 1 American and 2 German residents, the respective shares will be FR=0.25, US=0.25, DE=0.25+0.25.