Data Catalog

PATSTAT Global

Author:	European Patent Office
Distribution:	Users of PATSTAT Global
PATSTAT Edition:	2021 Autumn Edition
Version:	5.18
Date:	01.10.2021

Data Catalog for PATSTAT Global

The European Patent Office has, on behalf of the OECD Taskforce on Patent Statistics, prepared a database, called PATSTAT, that is designed to assist in statistical research into patent information.

Please direct enquiries about the database to patstat@epo.org. The discussion forum can be found in https://forums.epo.org/patstat .

Further information can be found on the PATSTAT information pages of the EPO web site at https://www.epo.org/patstat .

Document update history

Version	Date	Author	Explanation
0.3	31-05-2005	Heijna	First version
1.0	01-06-2005	Heijna	
2.01	23-06-2005	Heijna	
2.02	07-07-2005	Heijna	Improved relation types
2.1	15-07-2005	Heijna	Technical relations added, Citations re-modelled
2.11	20-07-2005	Heijna	Replenishment explained more clearly
2.2	25-07-2005	Heijna	Classes added
2.21	25-07-2005	Heijna	Improved business rules tables 001 and 010
2.22	26-07-2005	Heijna	Duplicate publication ID processing described, Primary keys highlighted, continuation type definition improved
2.23	02-08-2005	Heijna	IPC re-modelled, extracted citations definition improved
2.24	08-08-2005	Heijna	Various clarifications
3.0	09-09-2005	Rollinson	Enhancement of lay-out and format
		Heijna	Extension with physical model
3.1	13-09-2005	Heijna	Physical model refinements
3.11	22-09-2005	Heijna	Change in physical model
3.12	05-10-2005	Heijna	Textual changes and small improvements
3.13	07-10-2005	Heijna	Physical model diagram
3.2	21-11-2005	Heijna	Citation model upgraded, Conceptual model broken out
3.21	10-03-2006	Heijna	Comments added
3.22	20-04-2006	Rollinson	Foreword, Implementation rules, Directions for use added
3.23	08-09-2006	Heijna	Detailed improvements
3.24	02-05-2007	Rollinson	Updated “Directions for use”, Internet links (EPO site was rebuilt, many links changed)
3.25	27-08-2007	Rollinson	Added table for Classifications ICO-Y01N; US Applicant addresses ; improved country code coverage; added IPC advanced symbols; changed IPC symbol order to ‘as-in’ DOCDB ;
3.26	31-10-2007	Lingua	Revision of text
3.27	01-04-2008	Rollinson	Added table for DOCDB simple Family;
3.28	14-05-2008	Lingua	Small revision of text
3.29	23-09-2008	Rollinson/ Lingua	Added new table INPADOC family; redefined priorities with LMI=A and kind code W as equivalent to PARIS priorities.
3.30	15-10-2008	Lingua	Revision of text
3.31	16-04-2009	Lingua	Added new element PUBLN_FIRST_GRANT
3.32	22-05-2009	Lingua	Revision of text

4.00	17-06-2009	Rollinson	Major revision – replaced DOCDB database by DOCDB XML Exchange product as source. Removed access authorisation from element descriptions. Amended IPR_TYPE to include Design Patents, such as US with kind code 'S1'. References in the EPO Supplementary Search Reports to the original WO Search Report publication are now included as NPL citations.
4.06	14-07-2009	Rollinson/ Lingua	Text revised, updated diagrams, links, tables, element description
4.07	21-09-2009	Lingua	Adaptation to September 2009 edition
4.08	15-10-2009	Lingua	Renamed TLS217_APPLN_I_CLS with nanotech to TLS217_APPLN_ECLA as ECLA codes and all ICO codes (with nanotech) are included. Updated after Specs meeting with contractor 29th September.
4.09	18-02-2010	Lingua	Adaptation to April 2010 edition. Changes in loading IPC classes; reduction of artificial applications "D2" originating from to citations.
4.10	11-10-2010	Lingua	Routine for element PUBLN_FIRST_GRANT replaced by source from DOCDB XML, table TLS_221_INPADOC_PRS added, new citation kinds added in TLS_212_CITATION. Adaptation to September 2010 edition.
4.22	12-04-2011	Lingua	New permanent unique application identifier introduced in APPLN_ID, IPC CORE Level symbols no longer maintained by WIPO.
4.31	11-10-2011	Lingua	Adaptation to October 2011 edition, including creation of additional tables TLS222 and TLS223 for JP and US national classifications, adding elements CITN_GENER_AUTH and CITED_APPLN_ID in TLS212 and PUBLN_CLAIMS in TLS211
4.40	13-04-2012	Lingua	Adaptation to April 2012 edition
4.41	10-10-2012	Lingua	Adaptation to October 2012 edition
4.50	02-04-2013	Kracker	Complete revision of this document; Integration of physical model into the logical model; Deletion of table TLS217_APPLN_ECLA; Addition of TLS224_APPLN_CPC; Re-formulation of the domains of attributes in an DBMS-independent way. Various corrections and clarifications. Changes in data model (leading spaces removed from attributes APPLN_NR and PUBLN-NR)
5.00	14-10-2013	Kracker	Adaptation to October 2013 edition. Description of all elements of table TLS221_INPADOC_PRS. Inclusion of the extensions of PATSTAT Online. Minor clarifications and corrections;
5.01	01-04-2014	Kracker	Adaptation to 2014 Spring Edition; Various clarifications and minor corrections;

5.01.01	02-06-2014	Kracker	Correction of Diagram Section 3.2 “Logical Model Diagram”
5.02	15-10-2014	Kracker	Adaption to 2014 Autumn Edition: See section 7 “History of major changes to tables and attributes”. Clarification of stability of certain IDs: See section “4.3.2 Stable IDs”. Various other clarifications and minor corrections. Corrected errors in Logical Model Diagram (§3.2). Added chapter 8 “Known Deficiencies”
5.03	01-04-2015	Kracker	Adaption to 2015 Spring Edition: See section 7 “History of major changes to tables and attributes”. Pre-computed attributes which are only available in PATSAT Online: Their computation is specified in the new section “SQL scripts for computed attributes”. Various other clarifications and minor corrections;
5.04	01-10-2015	Kracker	Adaption to 2015 Autumn Edition: See section 7 “History of major changes to tables and attributes”. Section “SQL scripts for computed attributes” has been removed, because differences between the database schemas of PATSTAT Raw Data and PATSTAT Online data model have been removed.
5.05	01-12-2015	Kracker	Adaption to 2015 Spring – Amended Edition: See section 7 “History of major changes to tables and attributes”.
5.06	08-02-2016	Kracker	URLs to the EPO homepage have changed; Minor clarifications
5.07	01-04-2016	Kracker	Adaption to 2016 Spring Edition: See section 7 “History of major changes to tables and attributes”.
5.08	01-10-2016	Kracker	Adaption to 2016 Autumn Edition: See section 7 “History of major changes to tables and attributes”.
5.09	01-04-2016	Kracker	Adaption to 2017 Spring Edition: See section 7 “History of major changes to tables and attributes”.
5.10	01-10-2017	Kracker	Chapter 2 “Domain model” completely rewritten. Adaption to 2017 Autumn Edition: See section 7 “History of major changes to tables and attributes”.
5.11	01-04-2018	Kracker	Combined products “PATSTAT Biblio” and “PATSTAT Legal Status” into new product “PATSTAT Global”. Adaption to 2018 Spring Edition: See section 7 “History of major changes to tables and attributes”. Links to the EPO home page updated.
5.12	01-10-2018	Kracker	Adaption to 2018 Autumn Edition: See section 7 “History of major changes to tables and attributes”.
5.13	01-04-2019	Kracker	Adaption to 2019 Spring Edition: See section 7 “History of major changes to tables and attributes”.
5.14	01-10-2019	Kracker	Adaption to 2019 Autumn Edition: See section 7 “History of major changes to tables and attributes”.
5.15	01-04-2020	Kracker	Adaption to 2020 Spring Edition: See section 7 “History of major changes to tables and attributes”.
5.16	01-10-2020	Kracker	Adaption to 2020 Autumn Edition: See section 7 “History of major changes to tables and attributes”.

5.17	01-04-2021	Kracker	Adaption to 2021 Spring Edition: See section 7 “History of major changes to tables and attributes”.
5.18	01-10-2021	Hassler	Adaption to 2021 Autumn Edition: See Section 7 “History of major changes to tables and attributes”.

Table of contents

1 Introduction 14
2 Domain model 18
3 Logical model 24
4 Design principles 27
5 Table descriptions 41
6 Attribute description 88
7 History of major changes to tables and attributes 344
8 Known deficiencies 353

Detailed Table of contents

1 Introduction 14
1.1 About this document 14
1.2 About PATSTAT 14
1.3 PATSTAT product line 14
1.4 Data currentness and coverage 14
1.5 Data sources 15
1.5.1 Data source for bibliographic data 15
1.5.2 Data source for legal events 15
1.5.3 Data sources for person data 15
1.5.4 Data sources for harmonised names 16
1.5.5 Data sources for number of claims 16
1.6 Other databases for statistical purposes 16
1.7 Recommended reading on patent statistics 16
1.8 Correct citation of PATSTAT; copyright and trademark 16
1.9 Disclaimer 17
1.10 Help desk, discussion forum, feedback 17

2 Domain model 18
2.1 The 3 levels: Family – Application – Publication 18
2.2 Domain model diagrams 18
2.3 Description of domain objects 20
2.3.1 Abstract 20
2.3.2 Application 20
2.3.3 Citation 20
2.3.4 Classification 21
2.3.5 Family 21
2.3.6 Industry 22
2.3.7 Legal Event 22
2.3.8 Non-patent literature 22
2.3.9 Person 22
2.3.10 Publication 22
2.3.11 Technical field 23
2.3.12 Title 23

3 Logical model 24
3.1 Table naming convention 24
3.2 Logical model diagram 25

4 Design principles 27
4.1 Handling of double quotes and line breaks 27
4.2 Handling of missing or unknown values 27
4.3 Surrogate database keys 27
4.3.1 Pro and cons 27
4.3.2 Stable IDs 29
4.4 Application replenishment 30
4.4.1 Application replenishment for priorities 30
4.4.2 Application replenishment for citations 31
4.4.3 Allocating the APPLN_ID 32
4.5 Publication replenishment 33

4.5.1 Publications 33
4.5.2 Publication replenishment for citations 34
4.6 Relation types (Paris Convention priorities, continuations, etc.) 35
4.6.1 Rules 36
4.6.2 Continuation types 38

5 Table descriptions 41
5.1 TLS201_APPLN: Application 41
5.2 TLS202_APPLN_TITLE: Application title 44
5.3 TLS203_APPLN_ABSTR: Application abstract 45
5.4 TLS204_APPLN_PRIOR: Paris convention priority 46
5.5 TLS205_TECH_REL: Technical relation 47
5.6 TLS206_PERSON: Person 49
5.7 TLS207_PERS_APPLN: Link between Person and Application 52
5.8 TLS209_APPLN_IPC: International Patent Classification 53
5.9 TLS210_APPLN_N_CLS: National classification 55
5.10 TLS211_PAT_PUBLN: Patent publication 56
5.11 TLS212_CITATION: Citation 58
5.12 TLS214_NPL_PUBLN: Non patent literature publication 62
5.13 TLS215_CITN_CATEG: Citation category 65
5.14 TLS216_APPLN_CONTN: Application continuation 66
5.15 TLS222_APPLN_JP_CLASS: Japanese classification 67
5.16 TLS223_APPLN_DOCUS: US classification 68
5.17 TLS224_APPLN_CPC: Cooperative Patent Classification by application 69
5.18 TLS225_DOCDB_FAM_CPC: Cooperative Patent Classification by DOCDB family 70
5.19 TLS226_PERSON_ORIG: Unmodified person data 71
5.20 TLS227_PERS_PUBLN: Link between person and publication 74
5.21 TLS228_DOCDB_FAM_CITN: Citation between DOCDB families 75
5.22 TLS229_APPLN_NACE2: NACE2 industry classification 77
5.23 TLS230_APPLN_TECHN_FIELD: Classification by technical field 78
5.24 TLS231_INPADOC_LEGAL_EVENT: Legal event 79
5.25 TLS801_COUNTRY: Reference table of country codes 83
5.26 TLS803_LEGAL_EVENT_CODE: Reference table of legal event codes 84
5.27 TLS901_TECHN_FIELD_IPC: Mapping between technology fields and IPC 85
5.28 TLS902_IPC_NACE2: Mapping between IPC and industrial sectors 86
5.29 TLS904_NUTS: NUTS regional codes 87

6 Attribute description 88
6.1 Explanation of attribute description 88
6.2 ADDRESS_1, ADDRESS_2, ADDRESS_3, ADDRESS_4, ADDRESS_5 89
6.3 ADDRESS_FREEFORM 91
6.4 APPLN_ABSTRACT 93
6.5 APPLN_ABSTRACT_LG 94
6.6 APPLN_AUTH 95
6.7 APPLN_FILING_DATE 97
6.8 APPLN_FILING_YEAR 99
6.9 APPLN_ID 100
6.10 APPLN_KIND 102
6.11 APPLN_NR 104
6.12 APPLN_NR_EPODOC (deprecated) 106

6.13 APPLN_NR_ORIGINAL 108
6.14 APPLN_TITLE 109
6.15 APPLN_TITLE_LG 110
6.16 APPLT_SEQ_NR 111
6.17 CITED_APPLN_ID 113
6.18 CITED_DOCDB_FAMILY_ID 115
6.19 CITED_NPL_PUBLN_ID 116
6.20 CITED_PAT_PUBLN_ID 117
6.21 CITN_CATEG 120
6.22 CITN_GENER_AUTH 122
6.23 CITN_ID 124
6.24 CITN_ORIGIN 125
6.25 CITN_REPLENISHED 128
6.26 CITY 129
6.27 CLASS_SCHEME 130
6.28 CLASS_SYMBOL 131
6.29 CONTINENT 132
6.30 CONTN_TYPE 133
6.31 CPC_ACTION_DATE 135
6.32 CPC_CLASS_SYMBOL 136
6.33 CPC_DATA_SOURCE 138
6.34 CPC_GENER_AUTH 139
6.35 CPC_POSITION 140
6.36 CPC_STATUS 142
6.37 CPC_VALUE 143
6.38 CPC_VERSION 144
6.39 CTRY_CODE 145
6.40 DESIGNATED_STATES 146
6.41 DISCONTINUED 147
6.42 DOCDB_FAMILY_ID 148
6.43 DOCDB_FAMILY_SIZE 150
6.44 DOC_STD_NAME 151
6.45 DOC_STD_NAME_ID 153
6.46 DOCUS_CLASS_SYMBOL 154
6.47 EARLIEST_FILING_DATE 155
6.48 EARLIEST_FILING_ID 156
6.49 EARLIEST_FILING_YEAR 157
6.50 EARLIEST_PAT_PUBLN_ID 158
6.51 EARLIEST_PUBLN_DATE 159
6.52 EARLIEST_PUBLN_YEAR 160
6.53 EPO_MEMBER 161
6.54 EU_MEMBER 162
6.55 EVENT_AUTH 163
6.56 EVENT_CATEGORY_CODE 164
6.57 EVENT_CATEGORY_TITLE 165
6.58 EVENT_CODE 166
6.59 EVENT_DESCR 167
6.60 EVENT_DESCR_ORIG 168
6.61 EVENT_EFFECTIVE_DATE 169
6.62 EVENT_FILING_DATE 170

6.63 EVENT_ID 171
6.64 EVENT_IMPACT (deprecated) 172
6.65 EVENT_PUBLN_DATE 173
6.66 EVENT_SEQ_NR 174
6.67 EVENT_TEXT 175
6.68 EVENT_TYPE 176
6.69 EXTENSION_STATES 177
6.70 FEE_COUNTRY 178
6.71 FEE_PAYMENT_DATE 179
6.72 FEE_RENEWAL_YEAR 180
6.73 FEE_TEXT 182
6.74 FIRST_NAME 183
6.75 GRANTED 184
6.76 HAN_HARMONIZED 185
6.77 HAN_ID 186
6.78 HAN_NAME 187
6.79 INPADOC_FAMILY_ID 188
6.80 INT_PHASE 190
6.81 INTERNAT_APPLN_ID 192
6.82 INVT_SEQ_NR 194
6.83 IPC 196
6.84 IPC_CLASS_LEVEL 197
6.85 IPC_CLASS_SYMBOL 198
6.86 IPC_GENER_AUTH 200
6.87 IPC_MAINGROUP_SYMBOL 201
6.88 IPC_POSITION 202
6.89 IPC_VALUE 203
6.90 IPC_VERSION 204
6.91 IPR_TYPE 205
6.92 ISO_ALPHA3 206
6.93 JP_CLASS_SCHEME 207
6.94 JP_CLASS_SYMBOL 208
6.95 LAPSE_COUNTRY 209
6.96 LAPSE_DATE 210
6.97 LAPSE_TEXT 211
6.98 LAST_NAME 212
6.99 MIDDLE_NAME 213
6.100 NACE2_CODE 214
6.101 NACE2_DESCR 215
6.102 NACE2_WEIGHT 216
6.103 NAME_FREEFORM 217
6.104 NAT_CLASS_SYMBOL 220
6.105 NAT_PHASE 221
6.106 NB_APPLICANTS 222
6.107 NB_CITING_DOCDB_FAM 223
6.108 NB_INVENTORS 224
6.109 NOT_WITH_IPC 225
6.110 NPL_ABSTRACT_NR 226
6.111 NPL_AUTHOR 228
6.112 NPL_BIBLIO 230

6.113 NPL_CITN_SEQ_NR 232
6.114 NPL_DOI 233
6.115 NPL_EDITOR 235
6.116 NPL_ISBN 236
6.117 NPL_ISSN 238
6.118 NPL_ISSUE 241
6.119 NPL_PAGE_FIRST 243
6.120 NPL_PAGE_LAST 246
6.121 NPL_PUBLISHER 248
6.122 NPL_PUBLN_DATE 250
6.123 NPL_PUBLN_END_DATE 254
6.124 NPL_PUBLN_ID 256
6.125 NPL_TITLE1 257
6.126 NPL_TITLE2 259
6.127 NPL_TYPE 261
6.128 NPL_VOLUME 263
6.129 NUTS 266
6.130 NUTS_LABEL 267
6.131 NUTS_LEVEL 268
6.132 OECD_MEMBER 269
6.133 ONLINE_AVAILABILITY 270
6.134 ONLINE_CLASSIFICATION 271
6.135 ONLINE_SEARCH_DATE 273
6.136 ORGANISATION_FLAG 274
6.137 PARENT_APPLN_ID 275
6.138 PARTY_NEW 276
6.139 PARTY_OLD 277
6.140 PARTY_SEQ_NR 278
6.141 PARTY_TYPE 279
6.142 PAT_CITN_SEQ_NR 280
6.143 PAT_PUBLN_ID 281
6.144 PERSON_ADDRESS 282
6.145 PERSON_CTRY_CODE 284
6.146 PERSON_ID 287
6.147 PERSON_NAME 288
6.148 PERSON_NAME_ORIG_LG 290
6.149 PERSON_ORIG_ID 292
6.150 PRIOR_APPLN_ID 293
6.151 PRIOR_APPLN_SEQ_NR 294
6.152 PSN_ID 295
6.153 PSN_LEVEL 296
6.154 PSN_NAME 297
6.155 PSN_SECTOR 298
6.156 PUBLN_AUTH 299
6.157 PUBLN_CLAIMS 300
6.158 PUBLN_DATE 302
6.159 PUBLN_FIRST_GRANT 303
6.160 PUBLN_KIND 306
6.161 PUBLN_LG 307
6.162 PUBLN_NR 308

6.163 PUBLN_NR_ORIGINAL 309
6.164 RECEIVING_OFFICE 310
6.165 REF_DOC_AUTH 311
6.166 REF_DOC_DATE 312
6.167 REF_DOC_KIND 313
6.168 REF_DOC_NR 315
6.169 REF_DOC_TEXT 316
6.170 REG_PHASE 317
6.171 REINSTATE_COUNTRY 318
6.172 REINSTATE_DATE 319
6.173 REINSTATE_TEXT 320
6.174 RELEVANT_CLAIM 321
6.175 RESIDENCE_CTRY_CODE 322
6.176 ROLE 323
6.177 SOURCE 324
6.178 SOURCE_VERSION 325
6.179 SPC_EXTENSION_DATE 326
6.180 SPC_FILING_DATE 327
6.181 SPC_NR 328
6.182 SPC_PATENT_EXPIRY_DATE 329
6.183 SPC_TEXT 330
6.184 ST3_NAME 331
6.185 STATE 332
6.186 STREET 333
6.187 TECH_REL_APPLN_ID 334
6.188 TECHN_FIELD 336
6.189 TECHN_FIELD_NR 337
6.190 TECHN_SECTOR 338
6.191 UNLESS_WITH_IPC 339
6.192 WEIGHT 340
6.193 XP_NR 341
6.194 ZIP_CODE 343

7 History of major changes to tables and attributes 344

8 Known deficiencies 353

1 Introduction

1.1 About this document

This document describes the PATSTAT Global database. It contains diagrams showing the high-level structure, business rules, design principles, and detailed descriptions of the tables and attributes.

1.2 About PATSTAT

Patent statistics is used as an indicator of the inventive activity of companies or countries, and as an indicator of the patent system itself. As patent activity rose substantially over the last decades, the demand for patent data and statistics followed the same trend. It can help monitor the innovation activities, to better understand the innovation process and support policy decisions.

The EPO is an active member of the Patent Statistics Task Force led by the Organisation for Economic Co-operation and Development (OECD). Other members are the World Intellectual Property Organisation (WIPO), the Japanese Patent Office (JPO), the US Patent and Trademark Office (USPTO), Korean Intellectual Property Office (KIPO), the US National Science Foundation (NSF) and European Commission (EC). The EC is represented by Eurostat and by DG Research. Upon request of the Task Force, the EPO has created PATSTAT as the backbone data set for statistical analysis.

1.3 PATSTAT product line

PATSTAT consists of two individual products; the first product is covered in this document:

- **PATSTAT Global:**
  This is the core of PATSTAT (EPO product 14.24). It has a worldwide coverage and contains bibliographic information about patent applications and publications as well as legal event information.
- **PATSTAT EP Register:**
  It contains detailed bibliographic, procedural, and legal event information for EP patents (EPO product 14.24.1). This database is described in the PATSTAT EP Register Data Catalog.

1.4 Data currentness and coverage

This statistical database is a 'snapshot' of the source databases at a single point in time. Therefore, if you compare specific details with a register of a specific patent granting authority, you may detect some data differences, such as different names or dates; please consider that the authority may have corrected their databases in the meantime, i.e., since this statistical database 'snapshot' was created.

Typically, the date of data extraction from the source databases is end of January for the PATSTAT Spring Edition and end of July for the PATSTAT Autumn Edition.

The backfiles of the two main databases, i.e., DOCDB as the EPO worldwide bibliographic database and INPADOC as the EPO worldwide legal event database, are used to produce the corresponding edition containing all publications present in the databases at the end of week 05 (Spring edition) or week 30 (Autumn edition) for both DOCDB and INPADOC.

Information about the coverage of PATSTAT, DOCDB, and INPADOC data can be found

- as interactive chart:
  https://public.tableau.com/profile/patstat.support#!/vizhome/CoverageofPATSTAT2021SpringEdition/CoveragePATSTATGlobal
- as Excel files:
  https://www.epo.org/searching-for-patents/data/coverage/weekly.html

1.5 Data sources

1.5.1 Data source for bibliographic data

The main part of the data is extracted from DOCDB, the EPO worldwide bibliographic database. The backfile extraction format is called 'EXCHANGE FORMAT EPO - Patent Information Resource' and can be downloaded from the EPO website as DOCDB User Documentation[1].

1.5.2 Data source for legal events

For table TLS231_INPADOC_LEGAL_EVENT the data is extracted from the INPADOC worldwide legal event database.

1.5.3 Data sources for person data

For improved quality, names and addresses of applicants, owners, and inventors are taken from several sources which are described below.

- **EPO data**
  Person data for the EPO applications is taken from the EP Patent Register data.

- **USPTO data**
  The US data for names and addresses for published granted patents published after 1976-01-01 is taken from the USPTO's patent databases, as published weekly on USPTO web site.

  Starting with the publications of September 29th, 2005, we also take the names and address data for published applications from USPTO's patent database.

  The US data for names and addresses for patents published before 1976-01-01 (published grants) and September 25th, 2005 (published applications) is taken from EPO's DOCDB database.

[1] https://www.epo.org/searching-for-patents/technical/docdb.html , tab "Further information"

- **DOCDB**
  For all other applications person data is taken from DOCDB.

1.5.4 Data sources for harmonised names

There are several types of harmonised names available:

- DOCDB Standardised Name:
  This name is retrieved from DOCDB.
- PATSTAT Standardised Name (PSN):
  This effort for harmonising names and allocation of assignee sectors is done by ECOOM[2].
- OECD HAN:
  The Harmonised Applicant Name is computed by OECD[3].

1.5.5 Data sources for number of claims

Claims counts for EPO and US publications are provided as special data feeds directly from EPO and USPTO.

1.6 Other databases for statistical purposes

Almost all national and regional patent offices offer online Internet access to their registers. We recommend consulting them if you need to conduct in-depth research on a single national or regional patent granting authority.

The EPO offers a range of online databases for EP patent data that you may use for further analysis or verification of your findings. They can be found on the EPO homepage[4].

1.7 Recommended reading on patent statistics

For a thorough introduction to patent statistics, we recommend consulting the "OECD Patent Statistics Manual"[5].

1.8 Correct citation of PATSTAT; copyright and trademark

If you publish analyses based on this statistical database, please cite the source of the data including the name of the current version, e.g., “PATSTAT Global - 2021 Autumn Edition”.

The copyright to this database as distributed by the EPO remains with the EPO. "PATSTAT" is a registered trademark.

---

[2] KU LEUVEN; https://www.ecoom.be/en/data-collections/patstat-enhancements

[3] https://www.oecd.org/sti/inno/intellectual-property-statistics-and-analysis.htm

[4] https://www.epo.org/searching-for-patents/data/bulk-data-sets.html, https://www.epo.org/shop/en/Data-and-services/c/subscriptions

[5] https://www.oecd.org/science/inno/oecdpatentstatisticsmanual.htm

1.9 Disclaimer

The data in the PATSTAT databases is based on other EPO databases and on data provided to EPO on a voluntary basis by national and supranational patent authorities. EPO actively seeks to create and maintain a high-quality data basis for PATSTAT but cannot assume any legal liability or responsibility for the accuracy or completeness of the database[6].

In case legal certainty, accurate or complete data is needed, EPO strongly suggests contacting the responsible patent authorities.

EPO would appreciate if users of PATSTAT reported deficiencies to patstat@epo.org so that appropriate measures may be taken to correct them and improve PATSTAT.

1.10 Help desk, discussion forum, feedback

A help desk is available at patstat@epo.org. The discussion forum can be found in https://forums.epo.org/patstat .

Your feedback is very valuable and welcome. Please report any errors or suggestions for improvement to the help desk.

[6] Please note that we are not allowed to modify the data originating from non-EP documents and therefore cannot guarantee their correctness

2 Domain model

This section shows the relevant domain objects and their relationships in a graphical manner and describes each of them in more detail. The database tables which implement these domain objects are also given for easy reference to the PATSTAT logical model (section 3).

2.1 The 3 levels: Family – Application – Publication

The core domain object is the Application, which is a request for patent protection for an invention filed with the EPO or another patent office.

During the life of a patent, various publications are issued. An application has a least one publication, otherwise it would still be treated as confidential and would not be accessible in any database. Exceptions are applications that, for example, have been used as a priority or have been cited, but then revoked before publication. For reasons of consistency, these are kept as so-called “artificial applications” (see section 4.4 “Application replenishment” and Note A in section 3.2 “Logical model diagram”).

Applications which cover the same or similar invention are grouped into families. There exist several family definitions; consequently, there are several types of families. Each family contains one or more applications as family members. For each type of family, it applies that each application belongs to exactly one family of that type.

2.2 Domain model diagrams

The domain model is too complex to be visualized in a single diagram. So, this section contains multiple diagrams, each containing the central domain object Application and some related domain objects.

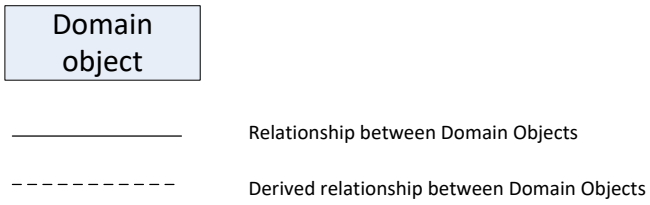


Figure 1: Explanation of symbols used

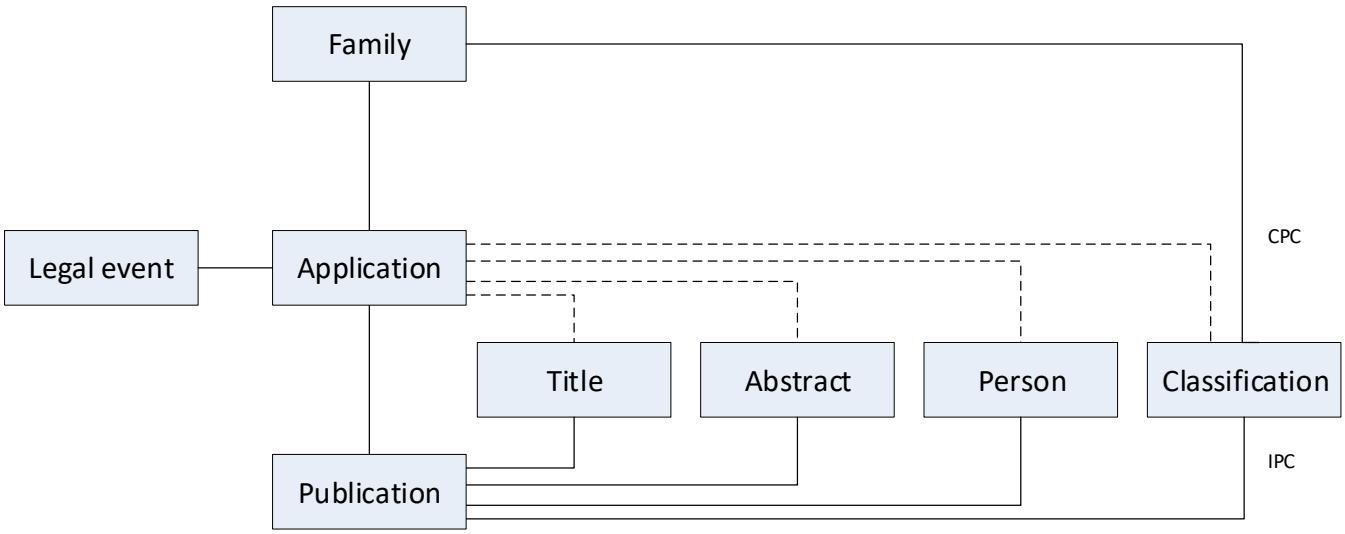


Figure 2: Main Domain Objects in the Family - Application - Publication hierarchy

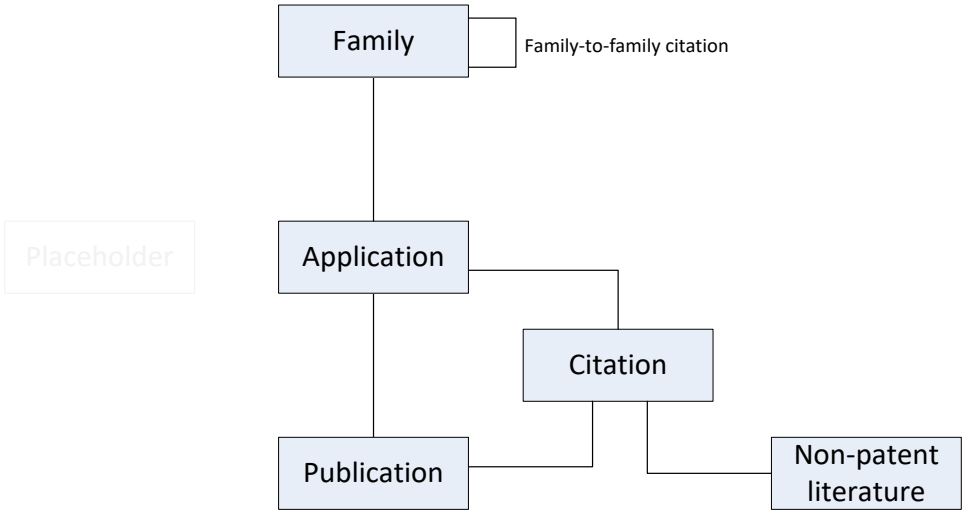


Figure 3: Citations in PATSTAT Global

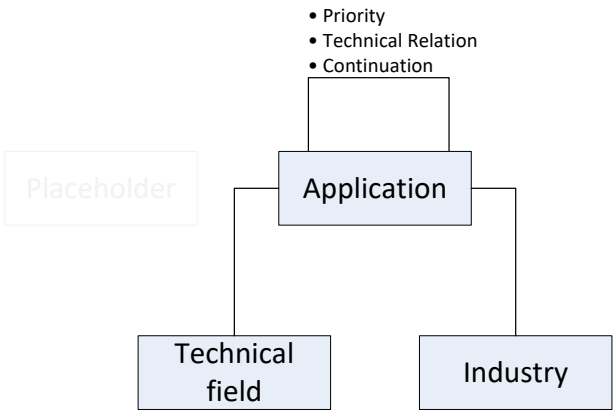


Figure 4: Further domain objects directly related to Application in PATSTAT Global

2.3 Description of domain objects

The domain objects are described in separate subsections, ordered alphabetically. References in brackets are made to the PATSTAT tables of the logical database model (section 3).

2.3.1 Abstract

This is the 1-paragraph summary of the invention which is shown on the first page of a publication. By design, in PATSTAT abstracts are related not to the individual publication, but to the application of the publication.

Abstracts can be in any language. PATSTAT contains only 1 abstract per application. Abstracts in English language are preferred. (TLS203_APPLN_ABSTRACT)

2.3.2 Application

The application is a request for patent protection of an invention. This is the central domain object. Most other domain objects are related to applications.

Every application (TLS201_APPLN) has at least 1 publication (TLS211_PAT_PUBLN). Every application belongs to exactly 1 simple family (also called DOCDB family) and to exactly 1 extended family (also called INPADOC family) (TLS201_APPLN).

Strictly speaking, title, abstract, persons and classifications are part of the publication. However, by design, in PATSTAT these domain objects are related not to the individual publication, but to the application of the publication.

Many more domain objects are directly related to applications. See these domain objects for details.

There are several optional relationships between applications:
- Priorities (TLS204_APPLN_PRIOR)
- Technical relations (TLS205_TECH_REL)
- Continuations (TLS216_APPLN_CONTN)

2.3.3 Citation

Citations (TLS212_CITATION) are references from patent publications to documents which are regarded as relevant for the patent procedure. They are identified in various stages in that procedure by various roles: by the applicant before application, during search and examination by the patent office, during an opposition procedure, by a third party etc.

Patent publications typically cite other patent publications or non-patent literature; in less frequent cases applications are also cited.

Each citation has one or more categories (TLS215_CITN_CATEG), which indicate the relevance of the citations. E. g. citation category “X” indicates that the claimed invention cannot be considered as novel due to the existence of the cited document.

For analysis, citations on the level of families (TLS228_DOCDB_FAM_CITN) are often regarded as more interesting than on the level of publications (c.f. section 2.1 The 3 levels: Family – Application – Publication).

2.3.4 Classification

Applications are classified according to their technical content by some symbol or code to facilitate searching. Multiple, hierarchically structured classification systems exist.

- IPC - International Patent Classification[7] (TLS_209_APPLN_IPC):
  This is maintained by WIPO[8] and used by all patent offices.
- CPC – Cooperative Patent Classification[9] (TLS224_APPLN_IPC):
  This in an extension of IPC. It has been created in 2013 and is maintained by EPO and the US patent office. Many major offices are nowadays using CPC, in addition to IPC.
  CPC symbols are assigned on a family level (TLS225_DOCDB_FAM_CPC). For ease of use, in PATSTAT they are also redundantly available on application level (TLS224_APPLN_IPC).
- FI (File Index) and F-Terms are used by the Japanese patent office for classification (TLS222_APPLN_JP_CLASS)
- USPC codes have been used by the US office for classification until recently (TLS223_APPLN_DOCUS)
- In the past, some offices have used their own national classification system (TLS210_APPLN_N_CLS)

2.3.5 Family

Applications which cover the same or similar invention are grouped into families. There are several definitions of families; consequently, there are several types of families. The EPO uses these types of families:

- Simple family, also called DOCDB family or Espacenet patent family:
  All applications which are member of the same simple family do have the same priorities. The technical content of these family members is regarded as (almost) identical, so their publications are sometimes called “equivalent”. (TLS201_APPLN)
- Extended family, also called INPADOC family:
  All applications which are member of the same extended family are directly or directly linked to the same root priority application. Usually, the applications are related to the same technical invention, but their individual content may differ. (TLS201_APPLN)

Every application belongs to exactly 1 simple family and to exactly 1 INPADOC family. The extended family is potentially “broader” than the simple family: Each extended family contains the applications of 1 or more simple families.

[7] http://www.wipo.int/classifications/ipc/en/
[8] http://www.wipo.int
[9] http://www.cooperativepatentclassification.org

2.3.6 Industry

The European Union uses NACE2 (Statistical Classification of Economic Activities in the European Community, version 2) to identify industries. Using a reference table based on IPCs, mostly NACE codes from the manufacturing industries are assigned to applications. (TLS229_APPLN_NACE2, TLS902_IPC_NACE2)

2.3.7 Legal Event

The Legal Event domain object represents procedural actions which change the (legal) status of an application or a granted patent. Some events are common to all jurisdictions, like refusal of an application, grant, entry into the national phase of a PCT application, payment of annual fee etc. (TLS231_INPADOC_LEGAL_EVENT, TLS803_LEGAL_EVENT_CODE)

2.3.8 Non-patent literature

Non-patent literature (NPL) (TLS214_NPL_PUBLN) can be cited by patent publications. NPLs can be any kind of public documents beside patent publications, such as, for example, books, articles in journals, databases, web pages, etc.
Certain NPLs, such as, for example, Derwent citations or Patent Abstracts of Japan, may themselves contain citations to patent publications.

2.3.9 Person

Persons may be legal persons (e.g., enterprises or any organisations) or natural persons (TLS206_PERSON, TLS226_PERSON). The domain object Person covers these roles:

- Applicants:
  The applicant is/are the person/s who filed the patent application. Depending on the phase of the application granting process, they may be also be the owner / proprietaries of the application or patent.
- Inventors:
  Only natural persons may be inventors.

An application may have at any point of time multiple applicants, inventors or representatives. These may also change over time. Only applicants are mandatory for an application. The same person can have multiple roles for the same application, like being applicant as well as inventor.

Persons are available as published by each publication (TLS227_PERS_PUBLN) as well as published by the most recent publication (TLS207_PERS_APPLN).

2.3.10 Publication

At defined stages in the application procedure, publications are issued (TLS211_PAT_PUBLN). There are several types of publications, each for a different purpose. Typically, the first application is published 18 months after its filing date or its priority date. The granted patent specification is published when patent protection has been granted. There are other kind of publications, e. g. corrections or publications of search reports, limitations etc.

Every application has at least 1 publication, because before the first publication an application is regarded as confidential and therefore not included in any public data set.

A patent publication typically consists of a front page, which contains meta-data (so called bibliographic data), the abstract and a representative image. On following pages are the detailed description of the invention, the claims, and the drawings.

2.3.11 Technical field

WIPO defined 35 technical fields which proved to be useful for statistical analysis. Using a reference table based on IPCs, these technical fields are assigned to applications. (TLS230_TECHN_FIELD, TLS901_TECHN_FIELD_IPC)

2.3.12 Title

This is the title of the invention which is shown on the first page of a publication. By design, in PATSTAT titles are related not to the individual publication, but to the application of the publication.

Titles can be in any language. PATSTAT Global contains only 1 title per application. Titles in English language are given preference over titles in other languages.
(TLS202_APPLN_TITLE)

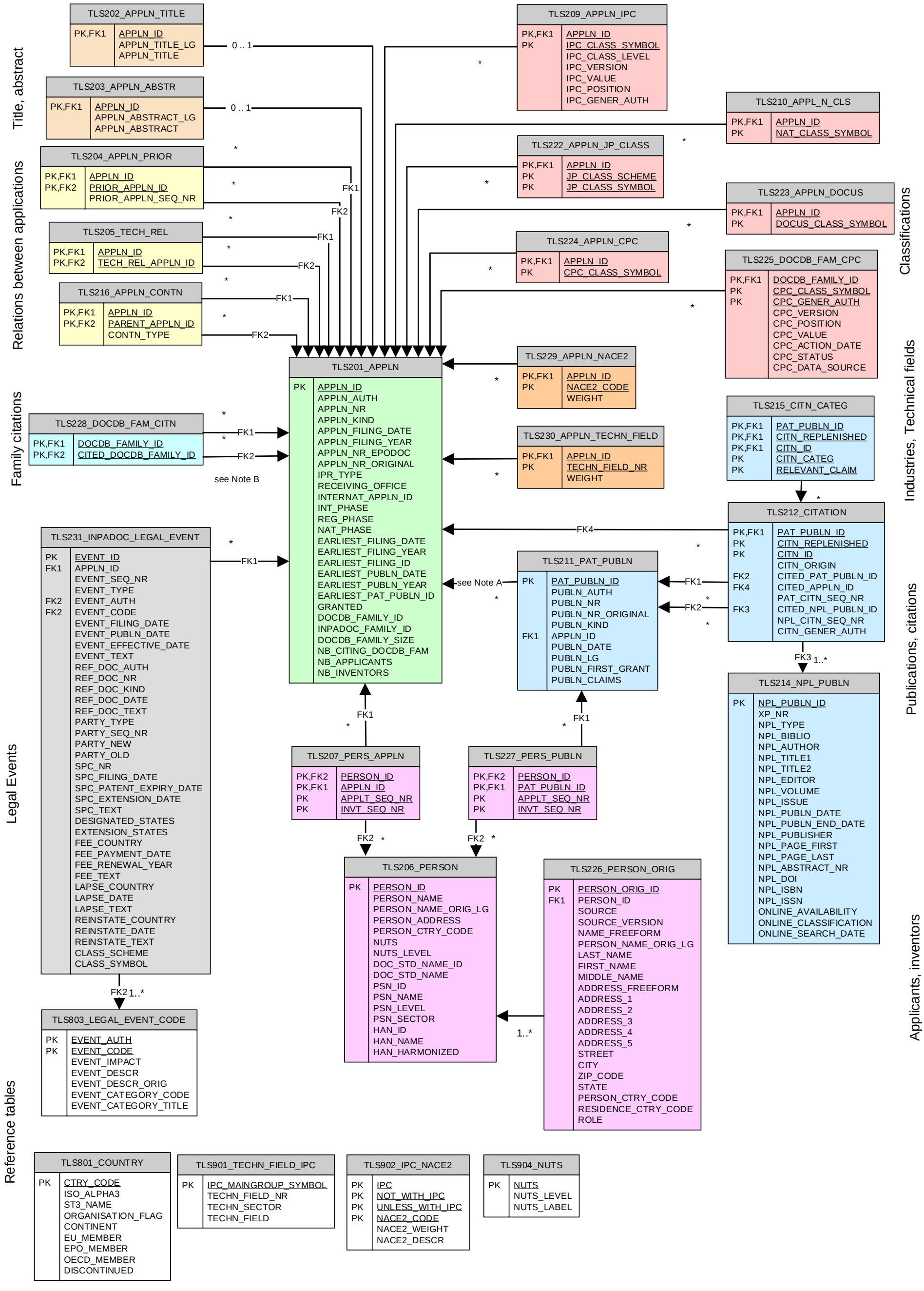
3 Logical model

3.1 Table naming convention

All table names in PATSTAT Global are of the form **TLSnnn_xxxx**

- TLS fixed prefix
- nnn unique number; the range of the number indicates it purpose
  200 range: data table
  800 range: reference table with data primarily managed by the EPO
  900 range: reference table with data primarily from external sources
- _ underscore
- xxxx one or more words indicating the content of the table

3.2 Logical model diagram



Legend:

0..1 cardinality: at most 1
1..n cardinality: at least 1
* cardinality 0 … n
PK This attribute is (part of) the Primary Key
FKn This attribute is (part of) the Foreign Key FK n

Underlined attributes constitute the Primary Key.

Note A:
Depending on their number range, some applications will or will not have publications. See section 4.4 "Application replenishment".

Applications with APPLN_ID			Number of Publications	cf. Range in section 4.4
0	-	900 000 000	1 or more	1
900 000 001	-	930 000 000	0	2
930 000 001	-	960 000 000	1 or more	3
960 000 001	-	999 999 999	0	4

Note B:
Both attributes in table TLS228_DOCDB_FAM_CITN must be linked to the attribute DOCDB_FAMILY_ID (and not APPLN_ID) of table TLS201_APPLN.

4 Design principles

4.1 Handling of double quotes and line breaks

Double quotes (“) are consistently replaced by single quotes (‘) in the data. This makes importing the data file, where double quotes are used to delimit text but will not appear within a text, easier.

Line breaking sequences (CR, LF) within strings are replaced by " `\n` ". This sequence e.g. occurs in the abstract text (attribute APPLN_ABSTRACT of table TLS203_APPLN_ABSTR) or the bibliographic data of non-patent literature (attribute NPL_BIBLIO in table TLS214_NPL_PUBLN).

4.2 Handling of missing or unknown values

It is a fact that for several documents, usually old ones, we are missing data, e.g., filing dates. In relational databases missing values are usually represented by NULL values, but these are hard to exchange in csv format.

PATSTAT data does not contain any NULL values and in fact all attributes may be defined as NOT NULL. Depending on the data type / domain, PATSTAT represents missing values like this:

- Missing values in attributes of type **date** are represented as '9999-12-31'.
- Missing values in attributes of type **string** are represented as zero length strings (like "") or as fixed length strings containing spaces.
- Missing values in **numerical** attributes are represented as number zero.

It is important to understand that the date 9999-12-31 means 'unknown’ or ‘not applicable’ date.

So if you write a query for all patents published after 2008-01-01, you will get a far larger number than you expect - you will also get all the 9999-12-31 ones. So remember to write something like

```
where PUBLN_DATE > '2008-01-01' and PUBLN_DATE < '9999-12-31'
```

4.3 Surrogate database keys

4.3.1 Pro and cons

A database 'key' is a shorthand reference to an entity (e. g. a publication, an application or a person) in a database. They uniquely identify such an entity. The key is just a number, which is meaningless to the human user, because it does not correspond to any number in

the business world, like an application number or publication number. Therefore, these surrogate keys are also called *technical identifiers*.

In PATSTAT you can easily recognize these technical identifiers because their names end with a *_ID* (not to be confused with names ending with *_NR*). There are several such technical identifiers, such as:

- APPLN_ID
- CITED_PAT_PUBLN_ID
- CITN_ID
- DOC_STD_NAME_ID
- DOCDB_FAMILY_ID
- EVENT_ID
- INPADOC_FAMILY_ID
- INTERNAT_APPLN_ID
- NPL_PUBLN_ID
- PARENT_APPLN_ID
- PAT_PUBLN_ID
- PERSON_ID
- PRIOR_APPLN_ID
- TECH_REL_APPLN_ID

The **advantage** of a surrogate key is that it is generally easier and more efficient to identify entities or to join tables with a surrogate key than with business identifiers.
As an example: The PAT_PUBLN_ID 387735615 identifies the publication EP 1665991 A9 published on 2007-03-28. Using business identifiers, you will need 4 attributes to uniquely identify this application: PUBLN_AUTH = EP, PUBLN_NR=1665991, PUBLN_KIND= A9 and PUBLN_DATE = 2007-03-28. Also, from a technical point of view, the database can store and manage surrogate keys more efficiently, which results in smaller database sizes and faster queries.

The **disadvantages** of a surrogate key are twofold:

- The identifier itself does not have any business meaning. You will always have to retrieve additional business data to be able to understand your result.
- The identifier can and likely will change between two editions of PATSTAT. In this case the same identifier may identify different business objects, like publications, in different editions of PATSTAT. So, if you want to exchange data between different PATSTAT editions, be sure to align the data not via surrogate keys but via business attributes.

  There are a few notable exceptions: For example, since April 2011, the surrogate key for applications, the APPLN_ID, remains stable. Also, the PAT_PUBLN_ID, the surrogate key for publications, does not change between different PATSTAT editions. The list of stable IDs and more information can be found in the next section 4.3.2 “Stable IDs”.

4.3.2 Stable IDs

The following attributes will not change between different PATSTAT editions, although in rare cases there may be exceptions. Therefore, these attributes can be used to link data from different PATSTAT editions. Moreover, the attributes APPLN_ID, PAT_PUBLN_ID and DOCDB_FAMILY_ID are taken directly from DOCDB, so they can also be easily be linked to other databases which are based on DOCDB.

- APPLN_ID since April 2011 edition, within Range 1
  (= not replenished applications, i.e., APPLN_ID ≤ 900 000 000)
- DOCDB_FAMILY_ID
- PERSON_ID since Oct 2013 edition
- PERSON_ORIG_ID since Oct 2013 edition
- PAT_PUBLN_ID since 2014 Autumn Edition, within Range 1
  (= not replenished publications,
  i.e. PAT_PUBLN_ID ≤ 900 000 000)
- EVENT_ID from table TLS231_INPADOC_LEGAL_EVENT
  since its introduction (2018 Spring Edition)

Exceptions to this stability assurance:

- APPLN_ID (within Range 1):
  These values are taken from DOCDB, so all restrictions of DOCDB apply:
  In instances where an application-reference has been re-keyed – technically speaking – the value of the APPLN_ID will remain stable and unchanged. There may be situations however – particularly as a result of online intellectual intervention – where this cannot be guaranteed.

- PAT_PUBLN_ID (within Range 1):
  These values are taken from DOCDB. The value of the PAT_PUBLN_ID is guaranteed to be stable, including in any event where the publication-identifier is corrected – also when the publication kind code has been corrected.

- DOCDB_FAMILY_ID:
  These values are taken from DOCDB, so all restrictions of DOCDB apply:
  The family-identifier is unique within the database, once used it will never be re-used, but its value cannot be guaranteed to be stable.

- PERSON_ID, PERSON_ORIG_ID:
  These attributes are the unique keys of person tables TLS206_PERSON and TLS226_PERSON_ORIG. These tables are also supposed to have no duplicates in their non-key values.
  However, in exceptional cases, e. g. due to data cleaning, duplicates may occur. These duplicates might be removed in later PATSTAT editions. This will result in IDs which have been available in older editions, but not in newer editions. However, once used, these IDs will never be re-used.

4.4 Application replenishment

Artificial applications are added to PATSTAT to manage doubt about applications which have not been captured in the DOCDB database from which PATSTAT is built. There are several cases:

- Application replenishment for priorities
- Application replenishment from citations
  - Applications originating from cited publications
  - Applications originating from cited applications

4.4.1 Application replenishment for priorities

By “priority” we here mean not only “Paris Convention priority”, but also other types of priorities which link one application to a “prior” application. The various types of priorities are stored in separate tables:

- TLS201_APPLN
  An PCT application in its regional/national phase contains in its attribute INTERNAT_APPLN_ID the APPLN_ID of its original PCT application
- TLS204_APPLN_PRIOR
  contains Paris Convention priorities
- TLS205_TECH_REL
  contains links between technically equivalent applications
- TLS216_APPLN_CONTN
  contains various relations like continuations, divisional applications, ...

There are cases where an application is claimed as priority, but this application is not known to DOCDB. Then we nevertheless assume that this prior application does really exist, although for some reason it is not in DOCDB. Therefore, we will create an artificial prior application in PATSTAT.
This can mean for example that you might find an application in table TLS201_APPLN, but not in EPO's search engine Espacenet as an application. However, you will find it in Espacenet if you search for it as a priority document.
Typically, these artificial applications are applications which have been withdrawn or abandoned before publication, but which the applicant has used as a priority, or in America, for continuation.

In more technical terms: If in the DOCDB backfile the application which is claimed as a priority in priority-claims for `data-format="docdb"` has no corresponding application-reference in DOCDB, then an artificial application must be created.

The example shows how the attributes of these artificial applications are populated:
- APPLN_AUTH is set to `<country>US</country>` from `priority-claim`
- APPLN_NR is set to `sequence="1"` from `priority-claim` and
- APPLN_KIND is set to `<kind>A</kind>` from `priority-claim` ,
  all for `data-format="docdb"` .
- APPLN_FILING_DATE is set to `<date>20040802</date>` from `priority-claim`

- APPLN_ID: Allocate a unique value incrementally, starting at 900 000 001.

4.4.2 Application replenishment for citations

There are two categories of replenished applications originating from citations:
a) Applications originating from cited publications
b) Applications originating from cited applications

4.4.2.1 a) Applications originating from cited publications

There are cited publications for which there is no publication reference in DOCDB. This includes cited patents which were extracted from Non-Patent Literature NPL citations. In this case an artificial publication is created in PATSTAT (see section 4.5 Publication replenishment). And we also create a matching application (see this section), because every publication must be assigned to an application.

The following business rules are applied:
Check if the cited publication has a publication-reference in DOCDB. If not, then create an artificial publication and an artificial application. The attribution of the artificial application is:

- APPLN_AUTH identical to the PUBLN_AUTH of the cited publication.
- APPLN_NR identical to the PUBLN_NR of the cited publication.
- APPLN_KIND = 'D2'.
- APPLN_FILING_DATE = '9999-12-31'
- IPR_TYPE = 'PI'
- APPLN_ID: Allocate a unique value incrementally, starting at 930 000 001.

4.4.2.2 b) Applications originating from cited applications

There are cited *applications* (in contrast to case a) publications) for which there is no application reference in DOCDB. In this case an artificial application is created in PATSTAT.

The following business rules are applied:
Check if the cited application has an application-reference in DOCDB and if not, then create an artificial application. The attribution of the artificial application is:

- APPLN_AUTH identical to the APPLN_AUTH of the cited application
- APPLN_NR identical to the APPLN_NR of the cited application
- APPLN_KIND identical to the APPLN_KIND of the cited application; if not given then use "D3". Note that in 2018 Autumn Edition there was no occurrence of 'D3'.
- APPLN_FILING_DATE identical to the APPLN_FILING_DATE of the cited application, if not given then assign '9999-12-31':
  If the same artificial application is cited more than once and with different application filing dates, then the earliest application filing date will be replenished. (Note: This logic minimizes the replenishment with the default date ‘9999-12-31’).
- APPLN_ID: Allocate a unique value incrementally, starting at 960 000 001.

4.4.3 Allocating the APPLN_ID

When collecting all applications, priorities and cited documents for all of the publications registered in DOCDB, it is important to keep them separate.

Once all of the application-references, publication-references, priority-claims and cited-references have been collected, it is possible to start allocating the surrogate key values for APPLN_ID and PUBLN_ID for the artificial applications and publications.

Starting with the April 2011 edition, the DOCDB "doc-id" unique and stable identifier has been used to populate APPLN_ID for non-replenished applications instead of creating an own surrogate key. This attribute remains the same across PATSTAT editions and always refers to the same combination of application authority, application number and application kind. It has a numeric value of max. 9 digits.

There are for ranges of replenished artificial applications in PATSTAT.
(Note: range 1 is for non-replenished applications, also called non-artificial applications; these applications are registered in DOCDB and their APPLN_ID is below 900 000 000):

Range 2:
Artificial applications created in PATSTAT for priorities applications which are not registered in DOCDB. They are not stable across PATSTAT editions.
Range: 900 000 001 to 930 000 000

Range 3:
Artificial applications created in PATSTAT for applications originating from cited publications not recorded in DOCDB. They are not stable across PATSTAT editions.
Range 930 000 001 to 960 000 000, kind code "D2".

Range 4:
Artificial applications created in PATSTAT for applications originating from cited applications not recorded in DOCDB. They are not stable across PATSTAT editions.
Range 960 000 001 to 999 999 999, using the kind code "D3" if the citation given has no kind code.

PATSTAT edition	Number of applications in DOCDB with a subsequent publication	**Ranges of APPLN_ID for artificial applications**		
		Range 2: PATSTAT applications created from unpublished DOCDB priorities	**Range 3:** PATSTAT applications created from cited publications with no publication in DOCDB	**Range 4:** PATSTAT applications created from cited applications with no application in DOCDB
2021 Autumn	104 267 238	900 000 001 - 908 598 234	930 000 001 - 931 792 365	960 000 001 - 960 032 197
2021 Spring	100 833 747	900 000 001 - 908 486 456	930 000 001 - 931 767 920	960 000 001 - 960 028 820
2020 Autumn	98 544 271	900 000 001 - 908 407 629	930 000 001 - 931 776 478	960 000 001 - 960 026 499
2020 Spring	95 211 153	900 000 001 - 908 265 869	930 000 001 - 931 856 318	960 000 001 - 960 022 045
2019 Autumn	92 361 935	900 000 001 - 908 143 966	930 000 001 - 931 842 454	960 000 001 - 960 018 691
2019 Spring	89 367 614	900 000 001 - 908 033 848	930 000 001 - 931 878 653	960 000 001 - 960 018 477
2018 Autumn	86 854 991	900 000 001 - 907 926 079	930 000 001 - 931 862 546	960 000 001 - 960 017 747
2018 Spring	84 375 547	900 000 001 - 907 831 637	930 000 001 - 931 851 934	960 000 001 - 960 017 588
2017 Autumn	82 147 124	900 000 001 - 907 725 467	930 000 001 - 931 848 414	960 000 001 - 960 017 061
2017 Spring	79 973 618	900 000 001 - 907 675 433	930 000 001 - 931 840 960	960 000 001 - 960 015 826

2016 Autumn	78 351 200	900 000 001 - 907 615 814	930 000 001 - 931 830 595	960 000 001 - 960 015 376
2016 Spring	76 504 845	900 000 001 - 907 526 532	930 000 001 - 931 744 552	960 000 001 - 960 014 259
2015 Aut Amend	74 469 830	900 000 001 - 907 427 548	930 000 001 - 931 738 471	960 000 001 - 960 013 617
2015 Autumn	74 469 830	900 000 001 - 907 427 548	930 000 001 - 931 728 217	960 000 001 - 960 012 596
2015 Spring	72 642 820	900 000 001 - 907 335 048	930 000 001 - 931 721 184	960 000 001 - 960 013 317
2014 Autumn	71 081 761	900 000 001 - 907 230 282	930 000 001 - 931 700 084	960 000 001 - 960 013 768
2014 Spring	69 410 835	900 000 001 - 907 140 127	930 000 001 - 931 724 340	960 000 001 - 960 013 546
2013 Oct	67 766 435	900 000 001 - 907 099 488	930 000 001 - 931 714 237	960 000 001 - 960 014 115
2013 April	66 012 696	900 000 001 - 906 989 695	930 000 001 - 931 755 005	960 000 001 - 960 014 651
2012 Oct	64 571 194	900 000 001 - 906 913 465	907 000 001 - 908 677 881	909 000 001 - 909 014 510
2012 April	63 280 409	900 000 001 - 906 826 996	907 000 001 - 908 669 845	909 000 001 - 909 014 916
2011 Oct	61 570 794	900 000 001 - 906 561 807	907 000 001 - 908 550 321	909 000 001 - 909 010 181
2011 April	60 312 074	900 000 001 - 906 476 936	907 000 001 - 908 692 290	not applicable
2010 Oct	58 713 013	59 000 001 - 65 239 596	66 000 001 - 67 274 345	not applicable
2010 April	57 505 125	58 000 000 - 63 983 731	64 000 001 - 65 252 476	not applicable
2009 Sept	56 420 849	57 000 001 - 62 913 743	63 000 001 - 64 260 712	not applicable
2009 April	55 517 602	56 000 001 - 61 951 472	62 000 001 - 63 239 563	not applicable
2008 Sept	54 371 495	55 000 001 - 60 883 933	61 000 001 - 62 241 942	not applicable
2008 April	53 357 975	54 000 001 - 59 803 164	60 000 001 - 61 238 598	not applicable
2007 Oct	52 389 958	53 000 001 - 58 670 414	59 000 001 - 60 232 649	not applicable
2007 April	?	52 000 001 - 57 616 300	58 000 001 - 60 447 086	not applicable
2006 Sept	?	?	?	not applicable
2006 April	?	50 000 001 - 55 527 619	56 000 001 - 58 541 387	not applicable

Modification history		
Author of update	Date of update	Explanation of update
R. Heijna	20-07-2005	First version
J. Rollinson	June 2009	Source changed to XML DOCDB
D. Lingua	19-02-2010	Inserted comment on number of D2s
D. Lingua	31-03-2011	Changes due to introduction of "doc-id"
D. Lingua	04-08-2011	Introduction of cited applications
M. Kracker	15-03-2013	Introduction of fixed ranges; clarifications
M. Kracker	01-04-2014	Clarifications

4.5 Publication replenishment

4.5.1 Publications

The EPO maintains a database called DOCDB (also known as Patent Information Resource) covering over 90 countries. The database contains patent documents and utility model documents which have been published or laid open to public inspection. These documents are a 'snapshot' of the status of an application at various stages in the lifecycle of the processing of the application according to the law of the relevant Office. Typically the contents of the application are published sometime after the priority date, in the EPO this is 18 months. If a search report is available, it is published at the same time. At later

stages in the lifecycle, such as grant, the contents of the application are published again, possibly in amended form. The different publication events in the lifecycle of the processing of an application are distinguished by the system of Kind of Publication Codes as laid down in the publication “Kind code concordance list”[10].

4.5.2 Publication replenishment for citations

When a document is cited, it is checked whether this document is already in the database by comparing the patent authority (country), the document number and the document kind code.

However, in roughly 2% of the cited documents in table TLS212_CITATION there is no corresponding publication entry in the table of published documents TLS211_PAT_PUBLN. This means that we cannot be 100% certain which document is intended to be cited.

Even if a cited publication is not known to DOCDB, we assume this document does really exist because it has been cited. Therefore, in these cases we introduce artificial publications in table TLS211_PAT_PUBLN. The attribution of an artificial publication is:

- PUBLN_AUTH, PUBLN_NR and PUBLN_KIND are taken from the citation
- PUBLN_DATE is assigned '9999-12-31’, if no publication date is given.
- PUBLN_ID: Allocate a unique value incrementally, starting at 900 000 001. The PUBLN_ID number range 900 000 001 to 999 999 999 is exclusively reserved for artificial publications.

We also create artificial applications to match these artificial publications (see section 4.4.2.1a) Applications originating from cited publications (Range 3).

PATSTAT edition	**Number of Publications** in DOCDB	**Range of PAT_PUBLN_ID for artificial publications:** PATSTAT publications created from DOCDB cited publications with no publication in DOCDB
2021 Autumn	133 397 784	900 000 001 - 901 852 544
2021 Spring	129 0941 98	900 000 001 - 901 827 112
2020 Autumn	126 159 780	900 000 001 - 901 835 270
2020 Spring	121 434 899	900 000 001 - 901 924 997
2019 Autumn	119 608 795	900 000 001 - 901 910 726
2019 Spring	113 796 117	900 000 001 - 901 949 477
2018 Autumn	110 397 316	900 000 001 - 901 933 758
2018 Spring	107 239 083	900 000 001 - 901 922 610
2017 Autumn	104 283 526	900 000 001 - 901 918 639
2017 Spring	101 185 732	900 000 001 - 901 910 449
2016 Autumn	98 592 257	900 000 001 - 901 899 315
2016 Spring	96 044 918	900 000 001 - 901 805 460
2015 Autumn Amended	93 276 814	900 000 001 - 901 795 268
2015 Autumn	93 276 814	900 000 001 - 901 784 222
2015 Spring	90 812 863	900 000 001 - 901 775 950

[10] https://www.epo.org/searching-for-patents/data/coverage/regular.html for databases within the EPO in column "DOCDB"

2014 Autumn	88 725 979	900 000 001 - 901 752 404
2014 Spring	86 430 793	900 000 001 - 901 724 340
2013 Oct	84 019 544	900 000 001 - 901 714 237
2013 April	81 694 203	900 000 001 - 901 755 005
2012 Oct	80 883 905	81 000 001 - 82 677 881
2012 April	79 049 630	80 000 001 - 81 669 845
2011 Oct	76 817 848	77 000 001 - 78 550 321
2011 April	74 274 345	75 000 001 - 76 692 290
2010 Oct	72 887 199	73 000 001 - 74 274 345
2010 April	71 217 622	72 000 001 - 73 252 476
2009 Sept	69 711 942	70 000 001 - 71 260 712
2009 April	68 453 166	69 000 001 - 70 239 563
2008 Sept	66 946 928	67 000 001 - 68 241 942
2008 April	65 493 394	66 000 001 - 67 238 598
2007 Oct	64 132 954	65 000 001 - 66 232 649
2007 April	?	63 000 001 - 65 447 086
2006 Sept	?	61 000 001 - 63 541 387
2006 April	?	60 000 000 - 62 426 270

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	15-03-2013	First version

4.6 Relation types (Paris Convention priorities, continuations, etc.)

Applications may be linked in various ways to other applications. This section describes how to distinguish 6 cases of these link types. 4 out of these 6 cases are stored in PATSTAT.

Relation Type	**Explanation**	**stored in PATSTAT table**
case # 1	self-priority: ignored	--
case # 2	Paris Convention priority	TLS204_APPLN_PRIOR
case # 3	national/regional phase of international application	TLS201_APPLN
case # 4	change of IPR-type claimed	TLS216_APPLN_CONTN
case # 5	technical relation (see note below)	TLS205_TECH_REL
case # 6	domestic continuation (e. g. divisional, US continuations, …)	TLS216_APPLN_CONTN

Note for case # 5 (Technical relations):
Technically related documents are those patent documents whose technical content has been identified within the EPO as being considered equivalent. This relation is identified in

the EPO master documentation database DOCDB by setting the indicator priority-linkage-type, also known as Link Method Indicator LMI, to "T" for 'Technical'.
The "T" indicator has allowed extracting most of the technical relations in table TLS205_TECH_REL. However, due to the manual intervention needed to create technical relations, it is known that a certain number of technical relations, especially before 1990, do not have the indicator set to "T", thus appearing in PATSTAT as a Paris convention priority.

4.6.1 Rules

This section describes the rules to infer the relation type (also known as LMI or Link Method Indicator or Linkage Type) from DOCDB XML data.

Note that LMI (Link Method Indicator) is represented in DOCDB XML as `<priority-linkage-type>` element.
This decision tree is applied to `<priority-claim data-format="docdb">` sections of DOCDB:

Is the content of element `<document-id>` in `<priority-claim data-format="docdb">` identical to the content of element `<document-id>` in `<application-reference>` (differences in the `<date>` can be ignored)?			
	Y	case # 1	
	N	What is the value of `<priority-linkage-type>`?	
		it is missing	case # 2
		W, w	case # 3
		A, a	Is `<kind>` ='W'?
			Y: case # 2
			N: case # 4
		I	case # 2
		U	case # 4
		T	case # 5
		other	case # 6

Modification history		
Author of update	Date of update	Explanation of update
R. Heijna	03-05-2005	First version
R. Heijna	13-07-2005	Continuation type added
R. Heijna	22-07-2005	Table continuation types separated
J. Rollinson	19-09-2008	for LMI=A , APPL_KIND='W' changed to PRIO_KIND='W'
D. Lingua	08-10-2012	Added rule for LMI=A, but APPL_KIND='W'
M. Kracker	26-03-2013	Restructuring; no change in logic

M. Kracker	01-04-2019	Simplified rule for LMI=A, but APPL_KIND='W'
M. Kracker	01-04-2020	Case #4 is not ignored anymore

4.6.2 Continuation types

Note that LMI (Link Method Indicator) is represented in DOCDB XML as `<priority-linkage-type>` element.

This table lists several values of `<priority-linkage-type>` as they may occur for certain offices.

Section 4.6.1 "Rules" defines how these `<priority-linkage-type>` elements are mapped to relation types. This table goes one step further and defines for each case #4 (= change of IPR type) and #6 (= domestic continuation) the detailed continuation type, which is stored in attribute CONTN_TYPE of table TLS216_APPLN_CONTN (see the respective table and attribute description).

APPLN_AUTH of the priority	**LMI**	**Description**	**CONTN_TYPE**
WO (PCT)	0	Prior application claimed for an addition	ADD
WO (PCT)	1	Prior application claimed for continuation	CON
WO (PCT)	2	Prior application claimed for continuation in part	CIP
WO (PCT)	3	Prior application claimed for a division	DIV
AT	A	CITED APPLICATION CHANGED FROM PATENT TO UTILITY	P2U
AT	U	CITED APPLICATION CHANGED FROM UTILITY TO PATENT	U2P
AU	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
AU	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
BA	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
BR	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
BR	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
CA	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
CA	4	PRIOR APPLICATION CLAIMED FOR A DIVISION OFA DIVISION	DIV
CA	5	CLAIMED APPLICATION IS A SUPPLEMENTARY DISCLOSURE	SUP
CH	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
CH	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
CN	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
CN	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
CS	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
CZ	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
DE	1	DOMESTIC PRIOR CLAIMED FOR PATENT	INN
DE	2	DOMESTIC PRIORITY CLAIMED FOR UTILITY MODEL	INN
DE	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
DE	I	PATENT APPLICATION CLAIMED FOR UTILITY MODEL	-
DK	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
EP	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV

ES	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
FI	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
FR	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
GB	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
HK	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
HU	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
IE	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
IE	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
IE	C	COGNATE APPLICATION	CGT
IL	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
IL	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
IN	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
IN	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
IN	C	COGNATE APPLICATION	CGT
JP	1	DOMESTIC PRIORITY	INN
JP	A	CITED APPLICATION CHANGED FROM PATENT TO UTILITY	P2U
JP	U	CITED APPLICATION CHANGED FROM UTILITY TO PATENT	U2P
KR	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
KR	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
KR	A	CITED APPLICATION CHANGED FROM PATENT TO UTILITY	P2U
KR	U	CITED APPLICATION CHANGED FROM UTILITY TO PATENT	U2P
LU	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
LV	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
MX	A	CITED APPLICATION CHANGED FROM PATENT TO UTILITY	P2U
MX	U	CITED APPLICATION CHANGED FROM UTILITY TO PATENT	U2P
NL	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
NO	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
NZ	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
NZ	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
NZ	C	COGNATE APPLICATION	CGT
PH	1	PRIOR APPLICATION CLAIMED FOR A CONTINUATION	CON
PH	2	PRIOR APPLICATION CLAIMED FOR A CONTINUATION IN PART	CIP
PH	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
PL	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
RU, SU	6	DOMESTIC PRIORITY	INN
TW	0	PRIOR APPLICATION CLAIMED FOR AN ADDITION	ADD
US	1	PRIOR APPLICATION CLAIMED FOR A CONTINUATION	CON
US	2	PRIOR APPLICATION CLAIMED FOR A CONTINUATION IN PART	CIP

US	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
US	4	PRIOR APPLICATION CLAIMED FOR A SUBSTITUTE	SBS
US	5	CLAIMED APPLICATION IS ORIGINAL REISSUE SERIAL NUMBER	REI
US	B	ABANDONED APPLICATION CLAIMED FOR A CONTINUATION	CON
US	C	ABANDONED APPLICATION CLAIMED FOR A CONTINUATION IN PART	CIP
US	D	ABANDONED APPLICATION CLAIMED FOR A DIVISION	DIV
US	R	REQUEST FOR REEXAMINATION NUMBER	REI
YU	3	PRIOR APPLICATION CLAIMED FOR A DIVISION	DIV
YU	6	DOMESTIC PRIORITY	INN

Modification history		
Author of update	Date of update	Explanation of update
R. Heijna	22-07-2005	Table continuation types separated
R. Heijna	26-09-2005	Definition based on application authority
M. Kracker	27-06-2013	Table sort order changed
M. Kracker	01-10-2017	Entries for countries BR and CN added; Entry CA / 4 corrected
M. Kracker	01-04-2019	Entries for WO (PCT) added; Entry for NO LMI=C removed
M. Kracker	01-04-2020	Continuation types P2U and U2P added

5 Table descriptions

5.1 TLS201_APPLN: Application

This table contains the key bibliographical data elements relevant to identify the patent application. Most of the elements in this table can be found on the first page of a printed patent document. E.g.: application authority, application number and application filing date. From a database structure point of view, this table is very important because it links to many other database tables via the attribute APPLN_ID.

TLS201_APPLN	
Technical identifier (Primary key)	
APPLN_ID	
Business identifiers	
APPLN_AUTH	
APPLN_NR	
APPLN_KIND	
APPLN_FILING_DATE	
APPLN_FILING_YEAR	
APPLN_NR_EPODOC	
APPLN_NR_ORIGINAL	
IPR_TYPE	
RECEIVING_OFFICE	
Data about the route of the application	
INTERNAT_APPLN_ID	
INT_PHASE	
REG_PHASE	
NAT_PHASE	
Data from priorities	
EARLIEST_FILING_DATE	
EARLIEST_FILING_YEAR	
EARLIEST_FILING_ID	
Data from publications	
EARLIEST_PUBLN_DATE	
EARLIEST_PUBLN_YEAR	
EARLIEST_PAT_PUBLN_ID	
Data derived from publications	
GRANTED	

Family data	
DOCDB_FAMILY_ID	
INPADOC_FAMILY_ID	
DOCDB_FAMILY_SIZE	
NB_CITING_DOCDB_FAM	
Aggregated data	
NB_APPLICANTS	
NB_INVENTORS	

PRIMARY KEY	APPLN_ID
Alternate Key	APPLN_AUTH, APPLN_NR, APPLN_KIND, RECEIVING_OFFICE

Business rules	Only applications which have been published are included in PATSTAT. Exceptions are “artificial applications**” (**see section 4.4 "Application replenishment") which have been added to make the database consistent. **Known duplicates** Some applications are stored twice in DOCDB and therefore also in PATSTAT. You will have to consider this when you are counting applications. The rule of thumb is: • Duplicates on APPLN_AUTH and APPLN_NR - one APPLN_KIND 'A', the other APPLN_KIND 'T': Application identifiers refer to one and the same application. • Duplicates on APPLN_AUTH and APPLN_NR - one APPLN_KIND 'A', the other APPLN_KIND 'D': o for APPLN_AUTH(s) AT, AU, BG, NL and SE: Application identifiers refer to one and the same application o for all other APPLN_AUTH(s): Application identifiers refer to two separate applications • Duplicates on APPLN_AUTH and APPLN_NR - one APPLN_KIND 'A' and the APPLN_KIND 'K', 'L', 'M' or 'N' : Application identifiers refer to two separate applications As another rule of thumb: If in doubt which filing to consider for counting, count only those which have a publication.
Comments	A first filing, i.e. an application not claiming the priority of any other application, can be identified by its *absence* from table TLS204_APPLN_PRIOR. Also, attribute INTERNAT_APPLN_ID in table TLS201_APPLN must be 0 because a first filing must not be the result of a PCT applications.

Modification history		
Author of update	Date of update	Explanation of update
R. Heijna	07-09-2005	First version

R. Heijna	07-10-2005	Continuations broken out
J. Rollinson	02-07-2009	added comment
M. Kracker	26-03-2013	added business rule for known duplicates
M. Kracker	01-10-2013	added PATSTAT Online extension attributes
M. Kracker	01-10-2015	Renaming and re-ordering several attributes. The tables TLS218_DOCDB_FAM and TLS219_INPADOC_FAM have been integrated. The language attributes have been moved to TLS202_APPLN_TITLE and TLS203_APPLN_ABSTR. Several attributes which have been available only in PATSTAT Online are now available in PATSTAT data as well.
M. Kracker	01-04-2016	New attributes APPLN_NR_ORIGINAL, INT_PHASE, REG_PHASE and NAT_PHASE added.
M. Kracker	01-04-2018	New attribute RECEIVING_OFFICE added. This attribute is also part of the alternate key.
M. Kracker	01-04-2019	Clarifications in business rules

5.2 TLS202_APPLN_TITLE: Application title

This table contains the English title of the application when available. In case there is no English title available, another language will be used if available.

TLS202_APPLN_TITLE				
	APPLN_ID			
	APPLN_TITLE_LG			
	APPLN_TITLE			
	PRIMARY KEY	APPLN_ID		
	FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
Business rules		Multiple titles may be published for any application, but only one title will be stored in PATSTAT, according to these rules (first applicable rule is applied) 1. most recent (according to publication date) title in English 2. most recent title in language of publication 3. most recent title in any other language		
Comments		n/a		
Modification history				
Author of update		Date of update	Explanation of update	
R. Heijna		07-09-2005	First version	
M. Kracker		01-10-2015	Attribute APPLN_TITLE_LG moved from table TLS201_APPLN to here	

5.3 TLS203_APPLN_ABSTR: Application abstract

This table contains the English language abstract, if available. If there is no abstract in English, then it contains the most recent abstract in another language.

TLS203_APPLN_ABSTR				
	APPLN_ID			
	APPLN_ABSTRACT_LG			
	APPLN_ABSTRACT			
	PRIMARY KEY	APPLN_ID		
	FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)

Business rules	Multiple abstracts may be published for any application, but only one abstract will be stored in PATSTAT, according to these rules (first applicable rule is applied) 1. most recent (according to publication date) abstract in English 2. most recent abstract in language of publication 3. most recent abstract in any other language
Comments	Starting with the 2020 Autumn Edition Euro-PCT applications, which are published without abstracts, are replenished with abstracts of their corresponding PCT application.

Modification history

Author of update	Date of update	Explanation of update
R. Heijna	07-09-2005	First version
M. Kracker	01-10-2015	Attribute APPLN_ABSTRACT_LG moved from table TLS201_APPLN to here
M. Kracker	01-10-2020	Euro-PCTs are replenished with abstracts from PCTs

5.4 TLS204_APPLN_PRIOR: Paris convention priority

This table contains the Paris Convention priorities of an application.

TLS204_APPLN_PRIOR				
	APPLN_ID			
	PRIOR_APPLN_ID			
	PRIOR_APPLN_SEQ_NR			
	PRIMARY KEY	APPLN_ID, PRIOR_APPLN_ID		
	FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
	FOREIGN KEY	PRIOR_APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
Business rules		APPLN_ID refers to the claiming application; PRIOR_APPLN_ID refers to the application of which the priority is claimed. These two foreign keys (applications) should be different ones, i.e. there is no "self-priority". There is a n:m relationship; multiple priorities may be claimed by one application and one priority may be claimed by multiple applications. Only "pure" priorities i.e. those according to the Paris Convention and published with an INID-code in the 30-series (WIPO ST.9) are included in this table. The relevant case is **case # 2** from section 4.6 "Relation Types".		
Comments		PCT applications (published with an INID-code in the 80-series; cf. WIPO ST.9) are no Paris Convention priorities, so they are not included in this table. A first filing, i.e. an application not claiming the priority of any other application, can be identified by its *absence* from this table Also, attribute INTERNAT_APPLN_ID in table TLS201_APPLN must be 0 to exclude national / regional phases of PCT applications.		
Modification history				
Author of update	Date of update	Explanation of update		
R. Heijna	07-09-2005	First version		
J. Rollinson	07-07-2009	added comment		
M. Kracker	15-10-2014	changed comment		

5.5 TLS205_TECH_REL: Technical relation

Technical relations are "priority-like" relations between applications which have been detected by EPO examiners but which have not been published by a patent office. From a statistical point of view you should consider them equal to the priority and continuation relations established in TLS204_APPLN_PRIOR and in TLS216_APPLN_CONTN.

TLS205_TECH_REL				
	APPLN_ID			
	TECH_REL_APPLN_ID			
	PRIMARY KEY	APPLN_ID, TECH_REL_APPLN_ID		
	FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
	FOREIGN KEY	TECH_REL_APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
Business rules		The relevant case is **case # 5** from section 4.6 "Relation Types".		
Comments		Technical relations are symmetric. When an application A and an application B are technically related, then B is also technically related to B. Consequently, if this table has a record A – B, then it will also contain a record B – A. Technical relations are a technical solution to connect old applications, which do not have priority information, into families. These relations are entered when detected by examiners or the EPO bibliographic data experts and no other priority-like relation exists between the applications. A more detailed explanation is further down. Most technical priorities are from FR, US, GB and DE applications, where large old collections, also from before 1900, exist. From 2008 onwards – with the introduction of the patent family building business rules – technical linking is very much the exception. There can however be no guarantee of completeness. This relation is also not published by Patent Offices. You can consider these technical relations as a priority-like relationship.		

Modification history		
Author of update	Date of update	Explanation of update
R. Heijna	07-09-2005	First version
M. Kracker	15.10.2014	Updated comment
M. Kracker	15.10.2016	Updated comment
M. Kracker	01.10.2017	Detailed explanation added

Why technical relations are created:

A technical family is created manually on request when documents disclosing identical subject matter (i.e. having identical description and drawings) are not automatically grouped together because they do not claim the same priority or combination of priorities.

The reasons why applicants may decide not to claim a priority are of various kinds: in some cases, the 12-month period foreseen in the Paris Convention might have been exceeded; in other cases, there might be economic reasons (e.g. innovation subsidies based on patent filings); yet in other cases, it could be related to the different ways in which IP offices - based on their respective IP laws - deal with patent continuations, divisionals and additions.

Let's look at an example: The EPO received a request for technical merge of the documents GB2542582, GB2542583 and GB2542584. These three patent applications, concerning a stretcher arrangement, were filed on the same day by the same applicant. The description and figures are in all three cases identical, but the claims cover three different aspects of the same invention: a pivoting arrangement for body support panels using coaxial actuators, a stretcher wheel assembly with solenoid activated locking mechanism and a pivoting body support for a stretcher, respectively.

Since they were filed on the same day, they could not claim each other's priority and originally entered the EPO search collection as first filings. This was changed by the EPO expert in order to create the technical relations which we now see in the EPO databases.

5.6 TLS206_PERSON: Person

Table that contains the key data on applicants and inventors such as: the person name, the address and the country/territory of residence (which is not necessarily the nationality). Several types of names are available:
- The name as delivered by the offices
- The name in original language, possibly in a non-Latin character set
- The name as standardised by the EPO (DOCDB standardised name)
- The PATSTAT standardised name
- The name as standardised by the OECD (OECD Harmonised Applicant Name)

TLS206_PERSON	
PERSON_ID	
Name and address delivered by the offices	
PERSON_NAME	
PERSON_NAME_ORIG_LG	
PERSON_ADDRESS	
PERSON_CTRY_CODE	
Region	
NUTS	
NUTS_LEVEL	
DOCDB standardised name	
DOC_STD_NAME_ID	
DOC_STD_NAME	
PATSTAT standardised name	
PSN_ID	
PSN_NAME	
PSN_LEVEL	
PSN_SECTOR	
OECD HAN name	
HAN_ID	
HAN_NAME	
HAN_HARMONIZED	
PRIMARY KEY	PERSON_ID
Alternate Key	PERSON_NAME, PERSON_NAME_ORIG_LG, PERSON_ADDRESS, PERSON_CTRY_CODE

FOREIGN KEY		REFERENCES	

Business rules	One and the same person may be recorded in different places in the source files. For some applications the inventor and the applicant may be the same person. Also applicants/inventors may occur in multiple applications. Where the name, the original language name, the address and the country/territory of a person in different places in the source files are identical (by case insensitive comparison), they are stored in this table only once. It is very likely although not absolutely sure that one entry in this table represents one (and not more) person in real life. On the other hand it is quite possible that a single person is represented by multiple entries of this table due to variations in name or address or changes of name and address. Several name harmonisation efforts try to reduce this ambiguity.
Comments	Persons are the legal or physical persons that have a relation with the patent granting procedure. Currently included are 2 roles a person may have: applicants and inventors. The sources of the Person tables are explained in section 1.5.3 “Data sources for person data”.

Modification history

Author of update	Date of update	Explanation of update
R. Heijna	07-09-2005	First version
J. Rollinson	October 2008	corrected column order
J. Rollinson	07-07-2009	extended comments
D. Lingua	08-03-2010	Added comment on US person data
D. Lingua	11-10-2011	Mention of OECD Working Group deleted
M. Kracker	01-10-2013	De-duplication rules changed: Comparison is now case-insensitive; Records with empty fields will be de-duplicated, too. Added PATSTAT Online extension attributes; Added comments on "see applicant" values
M. Kracker	01-04-2014	Removed comments on "see applicant" values
M. Kracker	15-10-2014	Comment updated
M. Kracker	01-04-2015	Attribute DOC_STD_NAME_ID_REPLENISHED has been removed; Order of attributes changed; Comment changed. TLS906_PERSON introduced;
M. Kracker	01-04-2016	Former EEE-PPAT names are now called PATSTAT Standardised Names and have

		been renamed to PSN_... . Attribute HRM_L1 has been removed.
M. Kracker	01-10-2016	Attributes NUTS and NUTS_LEVEL added
M. Kracker	01-04-2019	Attribute PERSON_NAME_ORIG_LG added
M. Kracker	01-10-2019	Table TLS906_PERSON has been removed. Its content has been merged into table TLS206_PERSON.

5.7 TLS207_PERS_APPLN: Link between Person and Application

This table links the applicants and inventors of the most recent publication to an application. Publications which contain only persons with non-Latin names (e.g. with Chinese characters) are not considered here.

TLS207_PERS_APPLN			
PERSON_ID			
APPLN_ID			
APPLT_SEQ_NR			
INVT_SEQ_NR			
PRIMARY KEY	PERSON_ID, APPLN_ID, APPLT_SEQ_NR, INVT_SEQ_NR		
FOREIGN KEY	PERSON_ID	REFERENCES	TLS206_PERSON (PERSON_ID)
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
Business rules	Table TLS207_PERS_APPLN links to the most recent[11] publication (if there is one) which contains at least a single applicant or inventor with a Latin name. Names which contain Latin as well as non-Latin characters are considered to constitute a Latin name. Some offices like CN, JP or KR frequently publish applicants and inventors with the non-Latin names. Only after several months (or never) the transliterated (now Latin) names are additionally made available. So right after the publication date table TLS207_PERS_APPLN might link to the persons of an earlier publication (if there is one), because the later publication does not have Latin names to link to.		
Comments	Conceptually, the combination of PERSON_ID and APPLN_ID should be unique. In practice, due to duplicates in the source data also the attributes APPLT_SEQ_NR and INVT_SEQ_NR must be part of the Primary Key.		

Modification history		
Author of update	Date of update	Explanation of update
R. Heijna	07-09-2005	First version
J. Rollinson	18-04-2006	Keys changed
M. Kracker	03-07-2013	Clarification of description; Primary Key changed; comment added
M. Kracker	01-10-2019	Publications with non-Latin person names only are ignored.

[11] Publications with an unknown publication date (PUBLN_DATE = '9999-12-31) are not considered here.

5.8 TLS209_APPLN_IPC: International Patent Classification

The table contains all international patent classifications linked to the applications. The set of classifications linked to a single application is a de-duplicated merge of all classifications of the various publication instances linked to the specific application. Additionally only the latest version of the IPC classifications is used. This means that the user does not have to worry about reclassifications because older applications will always be classified according to the latest IPC version.

TLS209_APPLN_IPC			
APPLN_ID			
IPC_CLASS_SYMBOL			
IPC_CLASS_LEVEL			
IPC_VERSION			
IPC_VALUE			
IPC_POSITION			
IPC_GENER_AUTH			
PRIMARY KEY	APPLN_ID, IPC_CLASS_SYMBOL		
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)

Business rules	In case the aggregation of the IPC symbols of all publications of a given application contains multiple IPCs with the same IPC_CLASS_SYMBOL, a 2 step de-duplication is performed: 1. For a given IPC symbol, only the highest IPC_CLASS_LEVEL is considered: Level A takes precedence over level C, and level C takes precedence over level S. 2. If there still exist multiple IPCs with the same IPC_CLASS_SYMBOL and IPC_CLASS_LEVEL, then the IPC from the latest publication takes precedence.
Comments	Information on classification according to the International Patent Classification (IPC) can be found in http://www.wipo.int/classifications/ipc/en/ . In principle the TLS209_APPLN_IPC table contains IPC 8 symbols according to the latest IPC 8 version. Therefore PATSTAT users do not need to worry about previous IPC editions IPC 1 to 7 or older versions of the IPC 8 edition when doing statistical analysis based on IPC codes. All these older symbols should in principle have been reclassified to the latest IPC 8 version. However, this is not always the case and these facts have to be kept in mind:

	1) DOCDB contains the MCD Master Classification Database. The MCD has IPC symbols allocated to over 90% of the documents in DOCDB; the remaining approx. 7% older documents are unlikely ever to be classified. 2) Some applications (about 580.000) have only IPCs of one of the old IPC editions 1 to 7 and have not been reclassified to IPC version 8. These applications are therefore not included in PATSTAT 3) Some applications will have an IPC 8 classification symbol that is not being used anymore in the latest IPC 8 version.

Modification history

Author of update	Date of update	Explanation of update
R. Heijna	07-09-2005	First version
R. Heijna	31-10-2005	IPC_GENER_AUTH added
R. Heijna	31-10-2005	generating office added
J. Rollinson	18-04-2006	Business rules and contents updated
J. Rollinson	Aug 2007	IPC Advanced added
J. Rollinson	Oct 2008	corrected column order
D. Lingua	02-07-2009	Text revised
D. Lingua	21-04-2010	Inserted warning on Core symbols
D. Lingua	05-04-2011	Adapted text due to Core discontinuation
M. Kracker	01-10-2013	Added PATSTAT Online extension attributes
M. Kracker	15-10-2014	Business rules clarified. Comment updated.
M. Kracker	10-10-2015	Removed pre-computed and redundant attributes IPC_SUBCLASS_SYMBOL and TECHN_FIELD_NR.
M. Kracker	01-04-2018	Comment revised

5.9 TLS210_APPLN_N_CLS: National classification

Some countries (GB, CH, CA, DE, FR, SE, ...) use national patent classification schemes beside the IPC. This table is a list of the national classifications linked to the respective national applications.

TLS210_APPLN_N_CLS			
APPLN_ID			
NAT_CLASS_SYMBOL			
PRIMARY KEY	APPLN_ID, NAT_CLASS_SYMBOL		
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)

Business rules	Classes can be present in DOCDB for all publication levels of an application. In PATSTAT these are re-grouped per application. Per application a national class symbol is present only once in PATSTAT unless the content of the source-field is unstructured in which case no de-duplication can be performed. Coverage is weak, particularly since US national classification migrated to table TLS223_APPLN_DOCUS in October 2011. Only a minority of applications in PATSTAT have a national class allotted.
Comments Rules	Dedicated tables TLS222_APPLN_JP_CLASS and TLS223_APPLN_DOCUS exist for JP and US national classification symbols.

Modification history

Author of update	Date of update	Explanation of update
R. Heijna	07-09-2005	First version
R. Heijna	10-03-2006	Business rules extended
J. Rollinson	07-07-2009	extended comment
D. Lingua	19-02-2010	Inserted comment
D. Lingua	04-08-2011	Eliminated national US and JP classification symbols from table

5.10 TLS211_PAT_PUBLN: Patent publication

This table contains the key bibliographical data elements relevant to identify patent publications. These elements can be found on the first page of printed patent documents. For example: publication authority, publication number, publication kind and publication date. This table is directly linked to the TLS201_APPLN table via the appln_id, a surrogate key that groups all the data elements from a single patent application. 2 important extra elements that cannot be found on a patent publication have been added:

a) The PUBLN_FIRST_GRANT:
indication that this publication was the first indication of a patent grant

b) PUBLN_CLAIMS:
number of claims (only available for a number of publishing authorities).

TLS211_PAT_PUBLN			
PAT_PUBLN_ID			
PUBLN_AUTH			
PUBLN_NR			
PUBLN_NR_ORIGINAL			
PUBLN_KIND			
APPLN_ID			
PUBLN_DATE			
PUBLN_LG			
PUBLN_FIRST_GRANT			
PUBLN_CLAIMS			
PRIMARY KEY	PAT_PUBLN_ID		
Alternate Key	PUBLN_AUTH, PUBLN_NR, PUBLN_KIND, PUBLN_DATE		
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
Business rules	This table also includes publications that are not present in DOCDB although they have been cited by another publication. In these cases the publications are created for the sake of data integrity (so-called ”artificial publications”). Only the key data of these artificial publications is known, such as authority, publication number and kind code, but not publication date, title or abstract or person data. For more information see section 4.5 “Publication replenishment”.		

Modification history		
Author of update	Date of update	Explanation of update
R. Heijna	07-09-2005	First version
R. Heijna	30-09-2005	Business rule changed (-> "last")
R. Heijna	20-04-2006	Business rule extended

D. Lingua	23-02-2009	PUBLN_FIRST_GRANT element added
J. Rollinson	2-07-2009	changed from DOCDB to DOCDB XML source
D. Lingua	04-08-2011	Element PUBLN_CLAIMS added
M. Kracker	01-04-2014	Correction of comment and foreign key: Multiple occurrences of publications with the same kind code are allowed.
M. Kracker	01-04-2016	New attribute PUBLN_NR_ORIGINAL.
M. Kracker	01-04-2018	Text of business rules revised

5.11 TLS212_CITATION: Citation

This table establishes the links between publications, applications and non-patent literature documents with regards to citations. Forward and backward citations are defined as well as the citation generating authority (e.g. search authority) and the procedural step in which the citation was created (e.g. search report or opposition procedure).

TLS212_CITATION			
PAT_PUBLN_ID			
CITN_REPLENISHED			
CITN_ID			
CITN_ORIGIN			
Patent literature			
CITED_PAT_PUBLN_ID			
CITED_APPLN_ID			
PAT_CITN_SEQ_NR			
Non-patent literature			
CITED_NPL_PUBLN_ID			
NPL_CITN_SEQ_NR			
Other attributes			
CITN_GENER_AUTH			
PRIMARY KEY	PAT_PUBLN_ID, CITN_REPLENISHED, CITN_ID		
FOREIGN KEY	PAT_PUBLN_ID	REFERENCES	TLS211_PAT_PUBLN (PAT_PUBLN_ID)
FOREIGN KEY	CITED_PAT_PUBLN_ID	REFERENCES	TLS211_PAT_PUBLN (PAT_PUBLN_ID)
FOREIGN KEY	CITED_APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
FOREIGN KEY	CITED_NPL_PUBLN_ID	REFERENCES	TLS214_NPL_PUBLN (NPL_PUBLN_ID)
Business rules	• PAT_PUBLN_ID refers to the *citing* publication. • CITED_PAT_PUBLN_ID refers to a *publication* being cited. • CITED_APPLN_ID refers to an *application* being cited. This cited application is *not* related to a publication cited by CITED_PAT_PUBLN_ID but is a valid citation on its own. • CITED_NPL_PUBLN_ID refers to a non-patent-literature being cited, which in turn may contain "hidden" references to patent publications.		

The two foreign keys for patent publications (PAT_PUBLN_ID and CITED_PAT_PUBLN_ID) should be different, i.e. there is no "self-citation".

Citations can represent a n:m relationship between publications: multiple publications may be cited in one publication and one publication may be cited by multiple others.

3 cases can be distinguished:

a) Patent citation, which is either a citation of a patent *publication* or patent *application*

b) Non-Patent Literature citation

c) Non-Patent Literature citation which refers to a patent *publication*

These 3 cases can be distinguished by the content of 5 attributes:

	a) Patent citation of a		b) NPL citation	c) NPL citation which refers to a patent publication
	patent publication	patent application		
CITED_PAT_PUBLN_ID	>0	=0	=0	>0
CITED_APPLN_ID	=0	>0	=0	=0
PAT_CITN_SEQ_NR	>0	>0	=0	=0
CITED_NPL_PUBLN_ID	=0	=0	>0	>0
NPL_CITN_SEQ_NR	=0	=0	>0	>0

Explanation of case c):
Patent publication citations *extracted from Non-Patent Literature* are also included, because the entry in table TLS212_CITATION will have the PAT_PUBLN_ID of the patent publication which is 'hidden' in the NPL citation stored in the column CITED_PAT_PUBLN_ID. For these citations the NPL_CITN_SEQ_NR of the relevant NPL-citations from which the patent citation was extracted is filled.

Comments

One publication can cite another publication or application multiple times, if the citation origin (see attribute CITN_ORIGIN) is different. E.g., the applicant and the examiner might cite the same publication, which would result in 2 records, with CITN_ORIG being “APP” resp. “EXA”. If you want to avoid double counting, make sure to count distinct citations only.

Regular information on citations is available to the EPO from the national patent authorities of the following countries: AP,

	AT, AU, BE, CZ, DE, EA, EP, ES, FR, GB, IT, JP, NL, SG, US, and WO. Batches of citations are present mainly for: BG, CH, DK, GR, KR, LU, and TR. For a complete and up-to-date coverage information see "Overview of citation data in REFI" (https://www.epo.org/searching-for-patents/data/coverage/regular.html) Before April 2016 due to a limitation in DOCDB, which is PATSTAT's source database, the number of citations was limited to 99 citations per publication and citation phase (CITN_ORIGIN), unless the citations were provided as "rich" citation. This affected the citations of less than 0.1% of the publications. Almost all of these missing citations were applicant citations of US publications.

Modification history

Author of update	Date of update	Explanation of update
R. Heijna	07-09-2005	First version
R. Heijna	13-09-2005	Alternate key added
R. Heijna	20-10-2005	Business rules extended Primary key redefined
R. Heijna	17-11-2005	Alternate key removed
R. Heijna	21-11-2005	Citation model upgraded
J. Rollinson	18-04-2006	Implementation rules added
D. Lingua	04-08-2011	Added elements CITED_APPLN_ID and CITN_GENER_AUTH
J Rollinson	30-04-2009	Patents hidden in NPL are now included in the NPL row in tls212_citation; they no longer have their own row.
J. Rollinson	02-07-2009	removed 5/6 as secret citations are not in DOCDB XML
J. Rollinson	02-07-2009	added WO reference
D. Lingua	13-07-2009	Updated citation information
D. Lingua	04-08-2011	Added elements CITED_APPLN_ID and CITN_GENER_AUTH
D. Lingua	18-04-2012	Updated comment information
M. Kracker	01-12-2013	Reordering of attributes. More detailed business rules
M. Kracker	15-10-2014	Comment added on how to avoid double counts of citations
M. Kracker	15-10-2015	Amended description of business rules
M. Kracker	01-12-2015	Extended *case c* description of business rules; added comment on DOCDBs 99 citations limit

M. Kracker	01-04-2016	Attribute NPL_PUBLN_ID renamed to CITED_NPL_PUBLN_ID. Business rules and comment amended.
M. Kracker	01-04-2018	New attribute CITN_REPLENISHED, which is also part of the Primary Key. Business rules amended.

5.12 TLS214_NPL_PUBLN: Non patent literature publication

This table contains bibliographical information on non-patent literature documents. The information is available as an unstructured string. In addition and to the degree possible, the information has been split up into multiple attributes.

TLS214_NPL_PUBLN										
	NPL_PUBLN_ID									
	XP_NR									
	NPL_TYPE									
	NPL_BIBLIO									
		Articles These attributes may only be populated if NPL_TYPE =					Online citations			
	NPL_AUTHOR	b	c	i		s	d	e	w	
	NPL_TITLE1	b	c	i		s	d	e	w	
	NPL_TITLE2	b	c	i	j	s	d		w	
	NPL_EDITOR	b								
	NPL_VOLUME	b	c	i	j	s	d		w	
	NPL_ISSUE		c	i	j	s	d		w	
	NPL_PUBLN_ DATE	b	c	i	j	s	d	e	w	
	NPL_PUBLN_END_DATE					s			w	
	NPL_PUBLISHER	b					d	e	w	
	NPL_PAGE_FIRST	b	c	i		s	d		w	
	NPL_PAGE_LAST	b				s			w	
	NPL_ABSTRACT_NR		c	i	j		d	e	w	
	NPL_DOI	b				s			w	
	NPL_ISBN	b			j	s			w	
	NPL_ISSN	b	c	i	j	s	d		w	
	ONLINE_AVAILABILITY							e	w	
	ONLINE_CLASSIFICATION						d			
	ONLINE_SEARCH_DATE								w	
	PRIMARY KEY	NPL_PUBLN_ID								
Business rules		From the 2016 Spring Edition onwards NPL citations which contain only strings like “none” or “See also references of WOxxxxxx” are removed to not distort citations counts.								
Comments		Attributes NPL_PUBLN_ID, XP_NR, NPL_TYPE and NPL_BIBLIO are always populated. All other attributes are may or may not be populated; some of them are sparsely populated. Depending on the NPL_TYPE many attributes have a (slightly) different meaning, which can be looked up in the attribute descriptions in chapter 6 “Attribute description”. Due to the way the EPO processes NPL citations, you may find slightly varying NPL data in other EPO								

	databases. Consequently, the data for the same NPL may vary from one PATSTAT edition to the next. E.g. there may be more, less or different attributes which have been populated or the data within an attribute may vary. The literature which is identified by this description is likely to be copyrighted.

The degree, to which an attribute is populated, highly depends on the NPL_TYPE (see section 6.127 “NPL_TYPE”). The values for NPL_TYPE are:

For poor NPL citations (no rich NPL structure):
- a Abstract citation of no specific kind

For articles:
- b Book citation
- c Chemical abstracts citation
- i Biological abstract citation
- j Patent Abstracts of Japan citation
- s Serial / Journal / Periodical citation

For online citations:
- d Derwent citation
- e Database citation
- w World Wide Web / Internet search citation

Actual population of fields in the 2021 Autumn Edition:

		Poor citations	Articles					Online		
Attribute		**a**	**b**	**c**	**i**	**j**	**s**	**d**	**e**	**w**
Number in 1 000		22085	782	35	1	30	7442	127	173	467
NPL_BIBLIO	in percent, rounded	100	100	100	100	100	100	100	100	100
NPL_AUTHOR			3	53	83		95	3	46	89
NPL_TITLE1			23	53	85		81	6	75	96
NPL_TITLE2			100	100	100	100	100			64
NPL_EDITOR			88							
NPL_VOLUME			8	71	70	99	59	91		38
NPL_ISSUE				68	20	99	36	91		34
NPL_PUBLN_DATE			95	72	61	93	71	5	59	96
NPL_PUBLN_END_DATE										1
NPL_PUBLISHER			69					100	88	
NPL_PAGE_FIRST			24				63			53
NPL_PAGE_LAST			14				56			49
NPL_ABSTRACT_NR				96	87	90		98	86	
NPL_DOI			1				11			20
NPL_ISBN			4				2			1
NPL_ISSN			2				14			23
ONLINE_AVAILABILITY									37	77
ONLINE_CLASSIFICATION								48		
ONLINE_SEARCH_DATE										81

Modification history		
Author of update	Date of update	Explanation of update
R. Heijna	07-09-2005	First version
R. Heijna	13-09-2005	Primary key redefined
R. Heijna	21-11-2005	Citation model upgraded
J. Rollinson	18-04-2006	Implementation rules added.
J. Rollinson	02-07-2009	added comment
M. Kracker	01-04-2016	New attribute NPL_TYPE. Business rule added.
M. Kracker	01-04-2017	18 new attributes added; Comment amended

5.13 TLS215_CITN_CATEG: Citation category

For most citations introduced during the search (citation origin is SEA), a citation category is added to the specific citation. Regular used citation categories are: X, Y and A. For example: category "X" is applicable where a document is such that when taken alone, a claimed invention cannot be considered novel or cannot be considered to involve an inventive step.

TLS215_CITN_CATEG				
	PAT_PUBLN_ID			
	CITN_REPLENISHED			
	CITN_ID			
	CITN_CATEG			
	RELEVANT_CLAIM			
	PRIMARY KEY	PAT_PUBLN_ID, CITN_REPLENISHED, CITN_ID, CITN_CATEG, RELEVANT_CLAIM		
	FOREIGN KEY	PAT_PUBLN_ID, CITN_REPLENISHED, CITN_ID	REFERENCES	TLS212_CITATION (PAT_PUBLN_ID, CITN_REPLENISHED, CITN_ID)
Business rules		The CITN_ID is a sequence number allocated to each citation made by a single document. The CITN_CATEG is the category of the citation as mentioned in search reports, e.g. X, Y, A, D, P. ….		
Comments				
Modification history				
Author of update		Date of update		Explanation of update
R. Heijna		07-09-2005		First version
R. Heijna		07-10-2005		Primary key definition
R. Heijna		21-11-2005		Citation model upgraded
J. Rollinson		18-04-2006		Implementation rules added.
J. Rollinson		2-07-2009		Added business rule
M. Kracker		01-10-2018		New attribute CITN_REPLENISHED, which is also part of the Primary Key.
M. Kracker		01-04-2019		New attribute RELEVANT_CLAIM, which is also part of the Primary Key.

5.14 TLS216_APPLN_CONTN: Application continuation

In a similar way as the TLS204_APPLN_PRIOR establishes the priority links between applications , the links between parent and child applications for various types relations such as: continuation (in part), divisional applications, internal priorities are defined via the TLS216_APPLN_CONTN table. Continuation (in part) is generally only applicable to US patent applications. This table should be considered as a priority-like relationship similar to the TLS204_APPLN_PRIOR table.

TLS216_APPLN_CONTN			
APPLN_ID			
PARENT_APPLN_ID			
CONTN_TYPE			
PRIMARY KEY	APPLN_ID, PARENT_APPLN_ID		
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
FOREIGN KEY	PARENT_APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)

Business rules	APPLN_ID refers to the continuation application; the PARENT_APPLN_ID refers to the application of which the APPLN_ID is a continuation. The two foreign keys (applications) should be different ones, i.e. there is no "self-continuation". There is a n:m relationship so a parent application may have multiple continuations and a continuation can have more than one parent. Only earlier applications for which a continuation is filed with the same authority (domestic) and for which the **continuation** is published with an INID-code in the 60-series (WIPO ST.9) are included in this table (plus inner priority, INID (23) as used by DE). The relevant case is **case # 6** from section 4.6 "Relation Types".
Comments	Continuations are e. g. divisional applications, additions, continuations in part, …

Modification history		
Author of update	Date of update	Explanation of update
R. Heijna	22-09-2005	First version

5.15 TLS222_APPLN_JP_CLASS: Japanese classification

FI and F-terms linked to JP application (only):

FI (File Index) has been developed to expand IPC in some technical fields. FI consists of an IPC symbol and an IPC-subdivision symbol and/or file discrimination symbol added to the IPC symbol.

F-TERMS (File Forming Terms) re-classify or further segment each specific technical field of IPC from a variety of viewpoints (i.e., objective, application, structure, material, manufacturing process, processing, etc.).

Japan's Patent Map Guidance System (PMGS) provides useful information about JP national classification of FI and F-terms in English. You may retrieve the classification list and an explanation for each classification:
https://www5.j-platpat.inpit.go.jp/pms/tokujitsu/pmgs_en/PMGS_EN_GM101_Top.action

TLS222_APPLN_JP_CLASS			
APPLN_ID			
JP_CLASS_SCHEME			
JP_CLASS_SYMBOL			
PRIMARY KEY	APPLN_ID, JP_CLASS_SCHEME, JP_CLASS_SYMBOL		
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
Business rules	The Japanese Classification schemes FI and FTERM, included in this table, are used by the Japanese Patent Office for carrying out patent application searches. The FI scheme is built on top of the International Patent Classification system (IPC), and is constantly being revised and updated. The FTERM scheme contains technical terms attributed from multiple viewpoints to facilitate computerised retrieval of patent documents. For more details see https://www.jpo.go.jp/e/system/patent/gaiyo/seido-bunrui/document/index/fi_f-term.pdf and https://www.jpo.go.jp/e/system/patent/gaiyo/seido-bunrui/#:~:text=FI%20%5BFile%20Index%5D%20and%20F,IPC%20%5BInternational%20Patent%20Classification%5D.		
Comments	n/a		
Modification history			
Author of update	Date of update	Explanation of update	
D. Lingua	04-08-2011	First version	
D. Lingua	11-10-2011	Updated link	
M. Kracker	01-04-2021	Updated link in business rules	

5.16 TLS223_APPLN_DOCUS: US classification

USPC codes linked to US applications (only):

US patent classification (USPC) codes are typically expressed in the following format "482/1" on US publications. The first number, 482, represents the class of invention. The number following the slash is the subclass of invention within the class. There are about 450 classes of invention and about 150 000 subclasses of invention in the USPC.

TLS223_APPLN_DOCUS			
APPLN_ID			
DOCUS_CLASS_SYMBOL			
PRIMARY KEY	APPLN_ID, DOCUS_CLASS_SYMBOL		
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)

Business rules	The classifications according to USPC (US patent classification) are used by the USPTO for carrying out patent application searches. The USPC system has a completely different structure than the International Patent Classification system (IPC). For more details follow this link http://www.uspto.gov/patents/resources/classification/help.jsp#5.
Comments	With the introduction of the Cooperative Patent Classification CPC on 2013-01-01 (http://www.cooperativepatentclassification.org/index.html) the DOCUS system has been superseded by the new CPC. The DOCUS coverage in PATSTAT has not been updated since then.
Implementation rules	n/a

Modification history

Author of update	Date of update	Explanation of update
D. Lingua	04-08-2011	First version
M. Kracker	01-10-2019	Updated description, business rules and comments

5.17 TLS224_APPLN_CPC: Cooperative Patent Classification by application

The table contains for each application its assigned cooperative patent classifications (CPC symbols). All applications of the same DOCDB family have the same CPC symbols assigned.

TLS224_APPLN_CPC			
APPLN_ID			
CPC_CLASS_SYMBOL			
PRIMARY KEY	APPLN_ID, CPC_CLASS_SYMBOL		
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)

Business rules	The data in this table is redundant and is derived from table TLS225_DOCDB_FAM_CPC. The table provides backward compatibility (for PATSTAT editions before Spring 2020) and improved ease of use.
Comments	See also table TLS225_DOCDB_FAM_CPC. Information on classification according to the Cooperative Patent Classification (CPC) can be found in https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/classification/cpc.html .

Modification history

Author of update	Date of update	Explanation of update
M. Kracker	07-03-2012	First version
M. Kracker	15-05-2012	Primary key extended by CPC_SCHEME
M. Kracker	01-10-2013	Added PATSTAT Online extension attribute
M. Kracker	15-10-2014	Business rules clarified. Comment updated.
M. Kracker	01-10-2015	Removed pre-computed and redundant attribute PC_MAINGROUP_SYMBOL.
M. Kracker	01-04-2019	Business Rules updated (invalid codes)
M. Kracker	01-04-2020	Complete restructuring and redefinition of this table. See now also table TLS225_DOCDB_FAM_CPC .

5.18 TLS225_DOCDB_FAM_CPC: Cooperative Patent Classification by DOCDB family

All applications of the same DOCDB simple family have the same cooperative patent classifications (CPC symbols) assigned. The same CPC symbol can be assigned to the same DOCDB family by one or more patent offices.
This table contains detailed information for each assignment.

TLS225_DOCDB_FAM_CPC				
	DOCDB_FAMILY_ID			
	CPC_CLASS_SYMBOL			
	CPC_GENER_AUTH			
	CPC_VERSION			
	CPC_POSITION			
	CPC_VALUE			
	CPC_ACTION_DATE			
	CPC_STATUS			
	CPC_DATA_SOURCE			
	PRIMARY KEY	DOCDB_FAM_ID, CPC_CLASS_SYMBOL, CPC_GENER_AUTH		
	FOREIGN KEY	DOCDB_FAMILY_ID	REFERENCES	TLS201_APPLN (DOCDB_FAMILY_ID)
Business rules		A small number of invalid or obsolete CPC classification codes can possibly occur. The EPO has preventive and corrective actions in place to avoid this as much as possible		
Comments		See also table TLS224_APPLN_CPC. Information on classification according to the Cooperative Patent Classification (CPC) can be found in https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/classification/cpc.html .		
Modification history				
Author of update		Date of update		Explanation of update
M. Kracker		01-04-2020		First version

5.19 TLS226_PERSON_ORIG: Unmodified person data

This table is best suited for detailed analysis of person data.
A row contains the name and address of a person (applicant and/or inventors; physical person or legal person). The data is taken from various data sources. It is kept in the "original" form, i.e. the data has not been cleaned, aggregated or otherwise modified. Depending on the data structure of each data source, not all attributes of this table are populated for every person.
Each row has one corresponding row in TLS206_PERSON. In these tables the data has been cleaned and unified and its table structure has been simplified and normalised.

TLS226_PERSON_ORIG	
Identifier and metadata attributes	
PERSON_ORIG_ID	
PERSON_ID	
SOURCE	
SOURCE_VERSION	
Name attributes	
NAME_FREEFORM	
PERSON_NAME_ORIG_LG	
LAST_NAME	
FIRST_NAME	
MIDDLE_NAME	
Address attributes	
ADDRESS_FREEFORM	
ADDRESS_1	
ADDRESS_2	
ADDRESS_3	
ADDRESS_4	
ADDRESS_5	
STREET	
CITY	
ZIP_CODE	
STATE	
PERSON_CTRY_CODE	
RESIDENCE_CTRY_CODE	
Other attributes	
ROLE	

PRIMARY KEY	PERSON_ORIG_ID		
FOREIGN KEY	PERSON_ID	REFERENCES	TLS206_PERSON (PERSON_ID)

Business rules	Rows where all attributes (except the primary key PERSON_ORIG_ID) are byte-wise identical are de-duplicated. The first table below explains which data source can populate which name and address attribute. The second table below details the data population of table TLS226_PERSON_ORIG for data from the DOCDB data source. DOCDB contains names of any combination of these three name formats: • Format “docdb**a**” contains **un**standardised names • Format “docdb” contains standardised names • Format “original” contains names in non-Latin characters
Comments	This table contains just name and address data. It cannot make reliable statements about persons in the real world. • It (quite likely) may be the case that 2 rows in the table represent one and the same person in the real world, due to variations of name or address data. • It may also be the (rare?) case that 2 persons in the real world are represented by the same row in this table, due to incomplete data.

	Data Source		
	DOCDB	**EP (Register)**	**USPTO Backfile; USPTO Frontfile DTD v4.2 - v4.5**
NAME_FREEFORM	✓	✓	
PERSON_NAME_ORIG_LG	✓		
LAST_NAME			✓
FIRST_NAME			✓
MIDDLE_NAME			✓
ADDRESS_FREEFORM	✓		
ADDRESS_1		✓	(✓)[12]
ADDRESS_2		✓	(✓)[4]
ADDRESS_3		✓	(✓)[4]
ADDRESS_4		✓	
ADDRESS_5		✓	
STREET			✓
CITY			✓

[12] There are only 1 710 USPTO addresses which use the fields ADDRESS_1, ADDRESS_2 or ADDRESS_3: they are all from USPTO Frontfile with DTD v4.2.

ZIP_CODE			✓
STATE			✓
PERSON_CTRY_CODE	✓	✓	✓
RESIDENCE_CTRY_CODE			✓ (only inventors)
ROLE			✓ (only applicants)

Which format exists on data source DOCDB?			**How isTLS226_PERSON_ORIG populated?**	
docdba (unstandardised)	**docdb (standardised)**	**original**	**NAME_FREEFORM**	**PERSON_NAME_ORIG_LG**
✓	✓	✓	docdba	original (see note 1 !!)
✓	✓		docdba	
✓		✓	docdba	original (see note 1 !!)
✓			docdba	
	✓	✓	docdb	original (see note 1 !!)
	✓		docdb	
		✓	original	original (see note 2 !!)

Note 1: only if different from value in NAME_FREEFORM;
otherwise PERSON_NAME_ORIG_LG stays empty.
Note 2: always

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	23-07-2013	First version
M. Kracker	01-10-2015	Data source for EP Register addresses are ADDRESS_1 to ADDRESS_5
M. Kracker	01-04-2016	Data source for USPTO addresses could also be ADDRESS_1 to ADDRESS_3.
M. Kracker	01-04-2019	Attribute PERSON_NAME_ORIG_LG added
M. Kracker	01-10-2019	Added explanation in Business Rules

5.20 TLS227_PERS_PUBLN: Link between person and publication

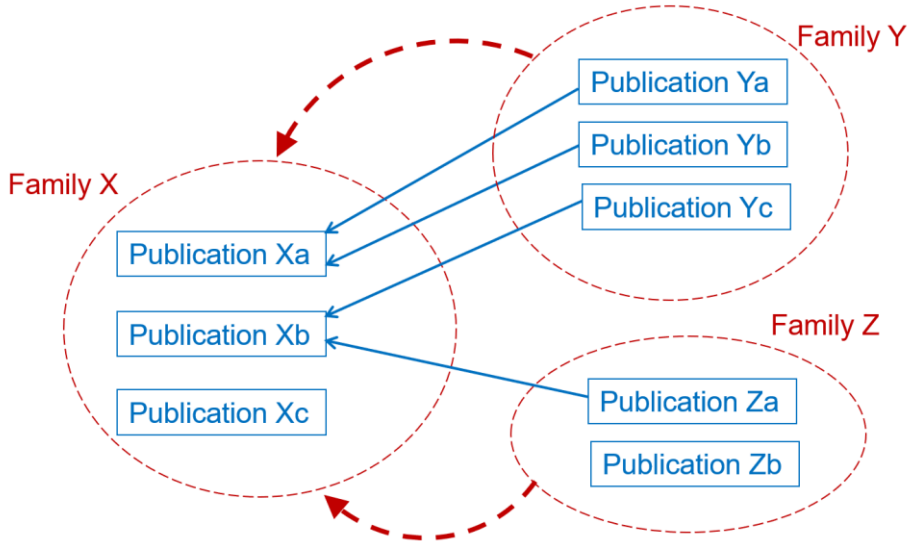
This table links each publication to its applicants and inventors. This can be used to analyse the changes of applicants / inventors at the times of their publication.

TLS227_PERS_PUBLN			
PERSON_ID			
PAT_PUBLN_ID			
APPLT_SEQ_NR			
INVT_SEQ_NR			
PRIMARY KEY	PERSON_ID, PAT_PUBLN_ID, APPLT_SEQ_NR, INVT_SEQ_NR		
FOREIGN KEY	PERSON_ID	REFERENCES	TLS206_PERSON (PERSON_ID)
FOREIGN KEY	PAT_PUBLN_ID	REFERENCES	TLS211_PAT_PUBLN (PAT_PUBLN_ID)
Comments	Conceptually, the combination of PERSON_ID and APPLN_ID should be unique. In practice, due to duplicates in the source data also the attributes APPLT_SEQ_NR and INVT_SEQ_NR must be part of the Primary Key.		

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	01-10-2013	First version

5.21 TLS228_DOCDB_FAM_CITN: Citation between DOCDB families

This table contains one entry for each pair of DOCDB simple families, where one member of a family cites at least one member of another family.
That means if multiple publications of one family cite one or multiple publication(s) / application(s) of another family, then this is counted as one citation between these 2 families.

TLS228_DOCDB_FAM_CITN			
DOCDB_FAMILY_ID			
CITED_DOCDB_FAMILY_ID			
PRIMARY KEY	DOCDB_FAMILY_ID, CITED_DOCDB_FAMILY_ID		
FOREIGN KEY	DOCDB_FAMILY_ID	REFERENCES	TLS201_APPLN (DOCDB_FAMILY_ID)
FOREIGN KEY	CITED_DOCDB_FAMILY_ID	REFERENCES	TLS201_APPLN (DOCDB_FAMILY_ID)
Business rules	Cited publications (the typical case) and well as cited applications (this is also possible) are considered when computing this table.		
Comments	To analyse inventions (represented by DOCDB families), one typically does not perform a low-level count by how many publications a publication is cited (c.f. blue solid arrows), but instead one counts by how many families a family is cited (c.f. red dashed arrows). In the example below, family X is cited by 2 other families: Family Y and family Z. 		

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	01-10-2013	First version

M. Kracker	01-04-2015	Formerly this table was called DOCDB_FAMILY_CITATION and was only available in PATSTAT Online. The order of the 2 columns has been reversed and one column has been renamed.
M. Kracker	01-10-2015	Foreign Keys now link to table TLS201_APPLN attribute DOCDB_FAMILY_ID
M. Kracker	01-04-2020	Explanation and diagram added

5.22 TLS229_APPLN_NACE2: NACE2 industry classification

This table tells to which degree an application belongs to one or more industries.

TLS229_APPLN_NACE2			
	APPLN_ID		
	NACE2_CODE		
	WEIGHT		

PRIMARY KEY	APPLN_ID, NACE2_CODE		
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
FOREIGN KEY	NACE2_CODE	REFERENCES	TLS902_IPC_NACE2 (NACE2_CODE)
Business rules	n/a		
Comments	This table is computed based on the reference table TLS902_IPC_NACE2 and the IPCs of an application. Consequently, applications without IPCs are not assigned to NACE2 codes. Note: The reference table TLS902_IPC_NACE2 maps IPCC codes to NACE codes which represent only manufacturing industries. Moreover, this table TLS229_APPLN_NACE2 includes all applications, even the ones whose applicants are universities, hospitals, and governmental organisations etc., which clearly are not manufacturers. Consequently – depending on your analysis – you may need to create your own mapping to NACE codes.		

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	01-04-2015	First version
M. Kracker	01-04-2017	Comment amended

5.23 TLS230_APPLN_TECHN_FIELD: Classification by technical field

This table tells to which degree an application belongs to one or more technical fields.

TLS230_APPLN_TECHN_FIELD			
	APPLN_ID		
	TECHN_FIELD_NR		
	WEIGHT		
PRIMARY KEY	APPLN_ID, TECHN_FIELD_NR		
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
FOREIGN KEY	TECHN_FIELD_NR	REFERENCES	TLS901_TECHN_FIELD_IPC (TECHN_FIELD_NR)
Business rules	n/a		
Comments	This table is computed based on the reference table TLS901_TECHN_FIELD_IPC and the IPCs of an application. Consequently, applications without IPCs are not assigned to technical fields.		

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	01-10-2015	First version

5.24 TLS231_INPADOC_LEGAL_EVENT: Legal event

This table holds the INPADOC data, which contains information on legal events that occurred during the life of a patent, either before or after grant. Typical events are request for examination, payment of renewal fees, lapse of the patent, change of ownership, withdrawal of the application, patent applications entering the national phase, patents which have been opposed or revoked, etc.

For EP patents this table contains

- most legal events which have been published in the EP Bulletin,
- legal events which have been delivered by the national offices during the national phase of the EP patent and
- post-grant events created by the EPO, namely payments, lapses and re-instantiations.

Events regarding priorities, application filings, grants (when linked to a specific publication) or publications are generally not included in this table. These events can, however, be derived from other PATSTAT tables.

TLS231_INPADOC_LEGAL_EVENT	
EVENT_ID	
APPLN_ID	
EVENT_SEQ_NR	
EVENT_TYPE	
EVENT_AUTH	
EVENT_CODE	
EVENT_FILING_DATE	
EVENT_PUBLN_DATE	
EVENT_EFFECTIVE_DATE	
EVENT_TEXT	
Reference to patent documents (application or publication)	
REF_DOC_AUTH	
REF_DOC_NR	
REF_DOC_KIND	
REF_DOC_DATE	
REF_DOC_TEXT	
Party	
PARTY_TYPE	
PARTY_SEQ_NR	
PARTY_NEW	

PARTY_OLD			
SPC (Supplementary Protection Certificate)			
SPC_NR			
SPC_FILING_DATE			
SPC_PATENT_EXPIRY_DATE			
SPC_EXTENSION_DATE			
SPC_TEXT			
Designated states and extension states			
DESIGNATED_STATES			
EXTENSION_STATES			
Fee payments			
FEE_COUNTRY			
FEE_PAYMENT_DATE			
FEE_RENEWAL_YEAR			
FEE_TEXT			
Lapses			
LAPSE_COUNTRY			
LAPSE_DATE			
LAPSE_TEXT			
Reinstatements			
REINSTATE_COUNTRY			
REINSTATE_DATE			
REINSTATE_TEXT			
Patent classification			
CLASS_SCHEME			
CLASS_SYMBOL			
PRIMARY KEY	EVENT_ID		
ALTERNATE KEY	APPLN_ID, EVENT_SEQ_NR		
FOREIGN KEY	APPLN_ID	REFERENCES	TLS201_APPLN (APPLN_ID)
FOREIGN KEY	EVENT_AUTH, EVENT_CODE	REFERENCES	TLS803_LEGAL_EVENT_CODE (EVENT_AUTH, EVENT_CODE)
Business rules	**Coverage of events:** Events regarding priorities, application filings, grants (when linked to a specific publication) or publications are generally *not* included in this table. These events can, however, be derived from other PATSTAT tables.		

<table>
<tr><td></td><td>
An entry with EVENT_CODE = "PGFP" (Post Grant Fees Paid) indicates that the annual renewal fee was paid in a specific country/territory in the national phase of a granted EP patent. Because this type of event typically repeats each year for each EP member state as long as it is valid in this member state, only the last PGFP event for each member state is recorded. Example: as soon as the 9th annual fee payment event for the FR national phase of an EP patent is recorded, the 8th annual fee payment event is removed from this table.

References to patent documents:

Event may reference documents (cf. attributes REF_DOC_xxx). No information is given whether this document is a patent application or a patent publication.

References to documents are given if PCT or EP patents are re-published by a regional / national office:
• PCT applications entering the regional / national phase are typically assigned a new regional / national number
• EP patents get a new national number by some offices (DE, AT, ES, EE, SK and GR)

If a legal event references a patent document it will be in one of 3 ways. Depending on the situation, certain attributes will always be populated (✓), some may be populated or not ([✓]). and some will always have the default value (empty cell):

<table>
<tr><th>Attribute</th><th>Case 1</th><th>Case 2</th><th>Case 3</th></tr>
<tr><td>REF_DOC_AUTH</td><td>✓</td><td></td><td></td></tr>
<tr><td>REF_DOC_NR</td><td>✓</td><td></td><td></td></tr>
<tr><td>REF_DOC_KIND</td><td>[✓]</td><td>✓</td><td></td></tr>
<tr><td>REF_DOC_DATE</td><td>[✓]</td><td>[✓]</td><td></td></tr>
<tr><td>REF_DOC_TEXT (i.e. free text)</td><td></td><td></td><td>✓</td></tr>
</table>
</td></tr>
<tr><td>Comments</td><td>
The source data used to create this table is described in the manual "Exchange Format EPO – Worldwide Legal Status" https://www.epo.org/searching-for-patents/data/bulk-data-sets/manuals.html , Download icon in section "INPADOC".

Linking legal events to other data

Via the attributes EVENT_AUTH and EVENT_CODE each legal event can be linked to an entry in table TLS803_LEGAL_EVENT_CODE. This reference table contains additional information about each type of legal event, e. g. a description, the group of the legal event or its impact on the life of the patent.
</td></tr>
</table>

	Via the attribute APPLN_ID each legal event can be linked to every PATSTAT table which contains an attribute APPLN_ID, e.g. to table TLS201_APPLN, which contains core data about each application. **Payments and patent validity** The payment of the annual renewal fee for an EP patent (see event code "PGFP" above) to an EPO member state is an extremely good indicator that this EP patent is valid or has been valid in that EPO member state.

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	01-04-2017	First version
M. Kracker	01-10-2017	New attribute EVENT_FILING_DATE
M. Kracker	01-04-2018	New attribute EVENT_ID, which is the Primary Key and which is stable across PATSTAT editions.

5.25 TLS801_COUNTRY: Reference table of country codes

Contains information about states/countries/territories and IP organisations, e.g. their two- and three-letter code, their (short) name and whether they are member of the EU, the EPO or the OECD. This table is based on WIPO standard ST.3.

TLS801_COUNTRY	
CTRY_CODE	
ISO_ALPHA3	
ST3_NAME	
STATE_INDICATOR	
CONTINENT	
EU_MEMBER	
EPO_MEMBER	
OECD_MEMBER	
DISCONTINUED	

PRIMARY KEY	CTRY_CODE
FOREIGN KEY	Via the CTRY_CODE attribute this table can be joined with any table which contains an attribute with two-letter codes of states/countries/territories and intergovernmental organisations.
Business rules	n/a
Comments	This table is based on WIPO standard ST.3 with additional public information. It is manually maintained by the EPO.

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	01-04-2014	First version
M. Kracker	01-04-2015	Addition of column ISO_ALPHA3

5.26 TLS803_LEGAL_EVENT_CODE: Reference table of legal event codes

This table contains all legal event codes which are used in EPO's worldwide legal event database (also called INPADOC database). Similar legal event codes are grouped into legal event categories.

TLS803_LEGAL_EVENT_CODE	
Legal event codes	
EVENT_AUTH	
EVENT_CODE	
EVENT_IMPACT	deprecated
EVENT_DESCR	
EVENT_DESCR_ORIG	
Legal event categories	
EVENT_CATEGORY_CODE	
EVENT_CATEGORY_TITLE	

PRIMARY KEY	EVENT_AUTH, EVENT_CODE
Business rules	n/a
Comments	This table corresponds to the Excel file "Legal status codes" available at the page https://www.epo.org/searching-for-patents/data/coverage/weekly . Via the attributes EVENT_AUTH and EVENT_CODE this table can be easily linked to table TLS231_INPADOC_LEGAL_EVENT. Information on the event categories, which are based on WIPO's ST.27 standard, can be found in the "INPADOC classification scheme" in https://www.epo.org/searching-for-patents/data/bulk-data-sets/manuals.html .

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	01-04-2017	First version
M. Kracker	01-10-2018	Change of attributes for legal event categories which are now based on WIPO ST.27; change of source

5.27 TLS901_TECHN_FIELD_IPC: Mapping between technology fields and IPC

This is the reference table which contains the mapping between 35 technology fields and the much more detailed IPC classification. These technology fields allow for the easy grouping of applications based on technology. The same technology fields are used by EPO and WIPO for their statistics.

TLS901_TECHN_FIELD_IPC	
IPC_MAINGROUP_SYMBOL	
TECHN_FIELD_NR	
TECHN_SECTOR	
TECHN_FIELD	

PRIMARY KEY	IPC_MAINGROUP_SYMBOL		
FOREIGN KEY	Via the IPC_MAINGROUP_SYMBOL attribute this table can be joined with any table which contains a compatible IPC or CPC attribute		
FOREIGN KEY	TECHN_FIELD_NR	REFERENCES	TLS209_APPLN_IPC (TECHN_FIELD_NR)
Business rules	n/a		
Comments	The content of this table is derived from http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/xls/ipc_technology.xls . More information on this technology classification can be found in "Concept of a Technology Classification for Country Comparisons" by Ulrich Schmoch, July 2008; http://www.wipo.int/edocs/mdocs/classifications/en/ipc_ce_41/ipc_ce_41_5-annex1.pdf		

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	?	First version
M. Kracker	01-10-2015	Order of attributes changed

5.28 TLS902_IPC_NACE2: Mapping between IPC and industrial sectors

A reference table which contains the mapping between the IPC classification and the NACE2 codes for industrial sectors. The industrial sectors allow for the grouping of applications based on the industry.

TLS902_IPC_NACE2	
IPC	
NOT_WITH_IPC	
UNLESS_WITH_IPC	
NACE2_CODE	
NACE2_WEIGHT	
NACE2_DESCR	

PRIMARY KEY	IPC, NOT_WITH_IPC, UNLESS_WITH_IPC, NACE2_CODE
FOREIGN KEY	Via the attributes IPC, NOT_WITH_IPC and UNLESS_WITH_IPC this table can be matched with any table which contains a compatible IPC or CPC attribute
Business rules	n/a
Comments	NACE2 is the Statistical Classification of Economic Activities in the European Community, Rev. 2 (2008) (**N**omenclature statistique des **a**ctivités économiques dans la **C**ommunauté **e**uropéenne). It serves a similar purpose than the SIC (Standard Industrial Classification) and the NAICS (North American Industry Classification System). EUROSTAT in co-operation with KU Leven / Belgium has provided a concordance table between IPC and NACE2, on which table TLS902_IPC_NACE2 is based. The original data and the description of the methodology can be retrieved from the https://forums.epo.org/concordance-table-between-ipc-and-nace2-9756 . This concordance table maps IPC sub classes / IPC main groups to the first 2-4 digits of the hierarchical NACE code. Note that this mapping maps all IPC codes only to NACE codes which represent manufacturing industries.

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	01-04-2015	First version
M. Kracker	01-04-2016	Links in comment updated, because a version of the mapping became available.
M. Kracker	01-04-2021	Comment and links updated.

5.29 TLS904_NUTS: NUTS regional codes

NUTS (Nomenclature of Territorial Units for Statistics) is a European Union standard for referencing the subdivisions of countries for statistical purposes. This reference table contains the regions of the NUTS levels 0 - 3.

TLS904_NUTS			
NUTS			
NUTS_LEVEL			
NUTS_LABEL			
PRIMARY KEY	NUTS		
FOREIGN KEY	NUTS	REFERENCES	TLS206_PERSON (NUTS)
Business rules	n/a		
Comments	This table contains NUTS information of NUTS version 2013.		

Modification history		
Author of update	Date of update	Explanation of update
M. Kracker	01-10-2016	First version
M. Kracker	01-04-2018	Table completely restructured.

6 Attribute description

6.1 Explanation of attribute description

Descriptor	Content
Name	Commonly used name of the field, e.g. "Application number"
Also Known As	Alternative common names of the field, e.g. "Dossier number" in case of EP applications
Description	Explanatory description of the field, e.g. "Numeric part of the identification of the application"
Domain	Description of the domain of values. Depending on the database management system you will use to manage this database, the appropriate data types must be chosen (e. g. nchar, nvarchar, date, integer, …).
Default value	The default value from the domain of values, if applicable
Source database	Name of the database that contains the original data, e.g. "DOCDB".
Source field name	Name of the field in the source database, e.g. "APPLT_SEQ_NR". This section may also contain instructions for EPO's IT supplier on how to process the data.
Source sub-field identifier	If necessary: Additional information to identify the source data.
Comments	Any further comments as deemed necessary

Modification history		
Author of update	Date of update	Explanation of update
R. Heijna	03-11-2004	First version
D. Lingua	14-07-2009	Preferred caption, Actuality and Source codes deleted from the table
M. Kracker	15-03-2013	Domain description does not depend on a specific DBMS

6.2 ADDRESS_1, ADDRESS_2, ADDRESS_3, ADDRESS_4, ADDRESS_5

Name: Address line 1, Address line 2, Address line 3, Address line 4, Address line 5
Also Known As: address
Description: First / Second / Third / Forth / Fifth address line of a person
Domain: string up to 500 characters
Default value: empty
Source database: EP Register
Source field name

```
<parties>
      <applicants change-gazette-num="2000/29">
            <applicant app-type="applicant" designation="all" sequence="1">
                  <addressbook>
                        <name>Seidel, Helmut</name>
                        <address>
                              <address-1>Fliederstrasse 19</address-1>
                              <address-2>65396 Walluf</address-2>
                              <country>DE</country>
                        </address>
                  </addressbook>
                  <nationality>
                        <country/>
                  </nationality>
                  <residence>
                        <country/>
                  </residence>
            </applicant>
      </applicants>
      <inventors change-gazette-num="2000/29">
            <inventor sequence="01">
                  <addressbook>
                        <name>Franta, Georg</name>
                        <address>
                              <address-1>Ulrich-Rapp-Strasse 18</address-1>
                              <address-2>87634 Obergünzburg</address-2>
                              <country>DE</country>
                        </address>
                  </addressbook>
            </inventor>
            <inventor sequence="02">
                  <addressbook>
                        <name>Dojan, Viktor</name>
                        <address>
                              <address-1>Ludwig-Strecker-Strasse 5</address-1>
                              <address-2>55129 Mainz</address-2>
                              <country>DE</country>
                        </address>
                  </addressbook>
            </inventor>
      </inventors>
```

Comments
The postal code and the city typically are in the last address line which is populated with data.

In PATSTAT Online due to data privacy reasons, the PERSON_ADDRESS has been emptied for all persons who might be a natural person (e. g. all inventors, or where the SECTOR attribute contains “INDIVIDUAL” or “UNKNOWN” or is empty.)

Modification history
Author of update - Date of update - Explanation of update
M. Kracker – 2015-10-01 – First version

6.3 ADDRESS_FREEFORM

Name: Full address in a single string
Also Known As: n/a
Description: Contains the full address in case the address is not available in structured form, where street, city, zip code, … are in different fields.
Domain: Up to 1000 characters
Default value: empty string
Source database: DOCDB
Source field name

```
<inventors>
      <inventor sequence="1" data-format="docdb">
            <inventor-name>
                  <name>STACY N SMITH</name>
            </inventor-name>
            <residence>
                  <country>US</country>
            </residence>
      </inventor>
      <inventor sequence="1" data-format="docdba">
            <inventor-name>
                  <name>STACY N. SMITH</name>
            </inventor-name>
            <address>
                  <text>305 Cottonwood Lane, NC 27540 Holly Springs,
UNITED STATES OF AMERICA (USA)</text>
            </address>
      </inventor>
      <inventor sequence="1" data-format="original">
            <inventor-name>
                  <name>Stacy N. Smith</name>
            </inventor-name>
      </inventor>
</inventors>

<applicants>
      <applicant sequence="1" data-format="docdb">
            <applicant-name>
                  <name>ERICSSON INC</name>
            </applicant-name>
            <residence>
                  <country>US</country>
            </residence>
      </applicant>
      <applicant sequence="1" data-format="docdba">
            <applicant-name>
                  <name>ERICSSON INC.</name>
            </applicant-name>
            <address>
                  <text>7001 Development Drive, 27709-3969 Research
Triangle Park,UNITED STATES OF AMERICA (USA)</text>
            </address>
      </applicant>
      <applicant sequence="1" data-format="original">
            <applicant-name>
                  <name>Ericsson Inc.</name>
            </applicant-name>
      </applicant>
</applicants>
```

Source sub-field identifier
`data-format="docdba"`

Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version

6.4 APPLN_ABSTRACT

Name: Abstract of application
Also Known As: n/a
Description: Abstract of the application
Domain: Up to 12 000 characters
Default value: n/a
Source database: DOCDB
Source field name

```
<abstract lang="EN" data-format="docdba" abstractsource="
National Office">
      <p>There is provided a floating surgical cannula. A
method of forming a surgical cannula by inserting a floating
surgical cannula at a location in need of surgery is
provided.</p>
</abstract>
<abstract lang="FR" country=WO doc-number="2005000001"
kind="A2" date="20050106" data-format="docdba" abstractsource="
National Office">
      <p>L'invention concerne une canule chirurgicale
flottante. L'invention concerne également un procédé de
formation d'une canule chirurgicale qui consiste à introduire
une canule chirurgicale flottante dans une zone à opérer.</p>
</abstract>
```

Source sub-field identifier

```
data-format="docdba"
```

Comments

Only one of possibly multiple abstracts is stored. See description of table TLS203_APPLN_ABSTRACT for details.

The average size of abstracts is 854 characters; maximum size is 9992 (as of Oct 2013).

Modification history

Author of update - Date of update - Explanation of update
R. Heijna - 13-05-2005 - First version
R. Heijna - 26-09-2005 - Oldest -> youngest
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 04-08-2011 - Addition of the PAJs
D. Lingua - 26-04-2012 - Eliminate comment on PAJs
M. Kracker - 26-03-2013 - Move comments to TLS203_APPLN_ABSTR table description
M. Kracker - 01-10-2013 - Increase suggested domain

6.5 APPLN_ABSTRACT_LG

Name: Language of abstract of application
Also Known As: n/a
Description: Language of the abstract of the application selected for and loaded in PATSTAT
Domain: 2 ASCII characters, according to ISO language codes (ISO 639-1) plus these DOCDB-specific extensions for languages:

bs Bosnian
hr Croatian
me Montenegrin

or spaces

Default value: spaces
Source database: DOCDB
Source field name

```
<abstract lang="EN" data-format="docdb" abstractsource="
National Office">
      <p>There is provided a floating surgical cannula. A
method of forming a surgical cannula by inserting a floating
surgical cannula at a location in need of surgery is
provided.</p>
```

Source sub-field identifier

```
data-format="docdb"
```

Comments
Use the value of this attribute for the abstract stored in the table TLS203_APPLN_ABSTR.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 13-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker – 01-10-2018 –language codes “bs”, “hr” and “me” added

6.6 APPLN_AUTH

Name: Application Authority
Also Known As: Country, State
Description: The competent authority, which is the national, international or regional patent office responsible for the processing of the patent application.
Domain: Up to 2 ASCII characters (A-Z), according to WIPO ST.3 (plus minor additions)
Default value: n/a
Source database: DOCDB
Source field name
1) Source for the standard applications:

```
<application-reference data-format="docdb" is-representative="N">
      <document-id>
                  <country>DE</country>
                  <doc-number>10331291</doc-number>
                  <kind>A</kind>
                  <date>20030710</date>
```

2) For priorities in DOCDB for which there is no application registered in DOCDB, use the authority (country) of the priority:

```
<priority-claims>
      <priority-claim sequence="1" data-format="docdb">
            <document-id>
                  <country>AE</country>
                  <doc-number>4000</doc-number>
                  <kind>A</kind>
                  <date>20000529</date>
            </document-id>
            <priority-active-indicator>Y</priority-active-indicator>
```

3) For artificial applications which were created for all artificial publications which were themselves artificially created for those cited publications, where the cited publications are not registered in DOCDB as publications: use the authority (country) of the cited publication:

```
<references-cited>
      <citation srep-phase="SEA" sequence="1">
            <patcit>
                  <document-id>
                        <country>AM</country>
                        <doc-number>199</doc-number>
                        <kind>A</kind>
```

4) For artificial applications which were created for applications originating from cited applications not recorded in DOCDB: use the authority (country) as cited in the cited application:

```
<citation srep-phase="APP" sequence="1">
      <patcit num="2" dnum="US19420613452A" dnum-type="application number">
            <document-id>
                  <country>US</country>
                  <doc-number>19420613452</doc-number>
                  <kind>A</kind>
                  <date>00000000</date>
```

```
            </document-id>
      </patcit>
</citation>
```

Source sub-field identifier
data-format="docdb"
Comments
Not to be confused with *country of origin*, which is the country of the applicant.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 16-11-2004 - First version
R. Heijna - 22-04-2005 - Source extended
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 15-03-2013 - Added artificial cited applications
M. Kracker - 15-05-2013 - Added exception to domain
M. Kracker - 01-04-2015 - Added ‘RH’ to domain
M. Kracker - 01-10-2019 – Comment amended

6.7 APPLN_FILING_DATE

Name: Application filing date
Also Known As: Date of receipt
Description: Date on which the application was physically received at the Patent Authority
Domain: Date (up to 9999-12-31)
Default value: 9999-12-31 (meaning 'unknown')
Source database: DOCDB

Source field name
1) Standard applications:

```
<application-reference data-format="docdb" is-representative="N">
      <document-id>
                  <country>DE</country>
                  <doc-number>10331291</doc-number>
                  <kind>A</kind>
                  <date>20030710</date>
```

2) Artificial applications from priorities:

```
<priority-claims>
      <priority-claim sequence="1" data-format="docdb">
            <document-id>
                  <country>DE</country>
                  <doc-number>10331291</doc-number>
                  <kind>A</kind>
                  <date>20030710</date>
            </document-id>
            <priority-active-indicator>Y</priority-active-indicator>
```

We assume that all priorities are accurately recorded in DOCDB. If a priority reference does not appear as an application reference, then in PATSTAT we create an artificial application with the authority (country), number kind and date of the priority. See APPLN_ID for the rules for creating the APPLN_ID for these artificial applications.
See rules for processing PRIOR_APPLN_SEQ_NR.

3) Artificial applications from citations:

```
<references-cited>
      <citation srep-phase="SEA" sequence="1">
            <patcit>
                  <document-id>
                        <country>US</country>
                        <doc-number>3380531</doc-number>
                        <kind>A</kind>
```

We assume that all cited references are publications. If a cited reference does not appear as a publication-reference, then in PATSTAT we create an artificial publication. See rules in element PUBLN_NR. We also create an artificial application, using the same country and number as the artificial publication, but we give an APPLN_FILING_DATE of 9999-12-31 and an APPLN_KIND of 'D2'.

4) For artificial applications which were created for applications originating from cited applications not recorded in DOCDB: use the application filing date as cited in the cited application, if not given or invalid then assign '9999-12-31':

```
<citation srep-phase="APP" sequence="1">
      <patcit num="2" dnum="US19420613452A" dnum-type="application number">
            <document-id>
```

```
                    <country>US</country>
                    <doc-number>19420613452</doc-number>
                    <kind>A</kind>
                    <date>00000000</date>
                </document-id>
            </patcit>
</citation>
```

Source sub-field identifier

```
data-format="docdb"
```

Comments

The legal filing date i.e. the date on which the legal protection starts may differ from the physical filing date. In case of a Divisional Application for instance the legal filing date is the one valid for the parent application which is earlier. It can also be later, e.g. when certain formal requirements are fulfilled later than the physical filing.

For (very) old applications the application filing date might not be known. This depends on the time period and is much more likely for patents from the 18th and 19th century than from the sixties of last century. Nevertheless, in many of these cases the publication dates might be given. So when working with those older applications, you should consider also the publication date as a proxy for a missing application filing date.

Modification history

Author of update - Date of update - Explanation of update
R. Heijna - 06-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 15-03-2013 - Added artificial cited applications
M. Kracker - 01-10-2019 – Comment amended

6.8 APPLN_FILING_YEAR

Name: Year of the application filing date
Also Known As: n/a
Description:
Domain: 4 digits in the form yyyy (e. g. 2015)
Default value: n/a
Source database: PATSTAT
Source field name: Derived from attribute APPLN_FILING_DATE of table TLS201_APPLN:

Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2015 – Computation explained

6.9 APPLN_ID

Name: Application identification
Also Known As: n/a
Description: Surrogate key: Technical unique identifier without any business meaning
Domain: Number 0 … 999 999 999
Default value: n/a
Source database: DOCDB (range 1), PATSTAT (ranges 2, 3, 4)
Source field name:

For range 1 (see below for definition of ranges):

```
<application-reference is-representative="YES" doc-id="11607218" data-
format="docdb">
      <document-id>
                  <country>DE</country>
                  <doc-number>8909720</doc-number>
                  <kind>U</kind>
                  <date>19890812</date>
      </document-id>
```

This corresponds to the XPath `/legal-status-document/legal-event/@doc-id` in the INPADOC database (EPO worldwide legal event database).

For ranges 2, 3 and 4:
APPLN_ID is set as described in section 4.4 "Application replenishment".

Source sub-field identifier: n/a
Source codes:
For range 1:

```
<application-reference is-representative="YES" doc-id="11607218" data-
format="docdb">
```

Comments
Previous to the April 2011 edition, a sequential number unique for each unique combination of the elements in the candidate primary key was attributed. The actual number had no particular meaning and would change from one edition to the next.

Starting with the April 2011 edition, the DOCDB "doc-id" unique and stable identifier has been used for all applications found in DOCDB (but not the number ranges 2, 3 & 4 below) to populate APPLN_ID instead of creating a PATSTAT-edition-specific surrogate key. DOCDB attribute "doc-id" contains a stable and unique identifier that will allow for linking up a number of EPO bulk data products through the application in a reliable way.
However, in exceptional cases some values of APPLN_ID might change even in number range 1 (see below). For details see section 4.3.2 “Stable IDs”.

There are 4 ranges of APPLN_ID:
Range 1: 1 to 900 000 000.
This range covers the filed applications which have a related publication in DOCDB. This range 1 is unique but not sequential (there are gaps in the sequence due to loading techniques). This attribute remains the same across PATSTAT editions and always refers to the same combination of application authority, application number and application kind. In

case an application is corrected, i.e. the application number and/or kind are changed, then it gets a new APPLN_ID. This is the only reason why a set of data (e.g. person names, publications) can relate to different APPLN_IDs across PATSTAT editions.

Range 2: from 900 000 001 to 930 000 000.
This range covers the artificial applications which are created in PATSTAT for prior applications, claimed as priorities, which do not have an application-reference in DOCDB.
The actual numbers in range 2 have no particular meaning and will change from one edition to the next.

Range 3: 930 000 001 to 960 000 000.
This range 3 covers the artificial filing applications with kind code D2 which are created in PATSTAT for those artificial publications which are also created in PATSTAT because these *publications* are cited, but do not have a publication-reference in DOCDB.
The actual numbers in this range have no particular meaning and will change from one edition to the next.

Range 4: 960 000 001 to 999 999 999
This range 4 covers the artificial filing applications with kind code D3 which are created in PATSTAT because these *applications* are cited.
The actual numbers in this range have no particular meaning and will change from one edition to the next.

See also section 4.4 "Application replenishment".

Note: For reasons of database consistency, there must be a dummy application with an APPLN_ID value of 0.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 15-04-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 31-03-2011 - Introduction of DOCDB unique stable identifier "doc-id"
M. Kracker - 15-03-2013 - Introduction of Range 4

6.10 APPLN_KIND

Name: Kind of Application
Also Known As: n/a
Description: Specification of the kind of application
Domain: Up to 2 ASCII characters:

- A patent
- U utility model
- W PCT application (in the international phase)
- T used by some offices (e. g. AT, DE, DK, ES, GR, HR, PL, PT, SI, SM, TR) for applications which are "translations" of granted PCT or EP applications
- P provisional application (US only)
- F design patent
- V plant patent
- D2, D3 artificial applications (see section 4.4 "Application replenishment")
- Other "exotic" kind codes:
  See DOCDB User Documentation (https://www.epo.org/searching-for-patents/data/bulk-data-sets/manuals.html), section 22.1 " 'Exotic' Kind-codes"
- Due to bad data, some artificial applications with an APPLN_ID > 960 000 000 have other values for APPLN_KIND

Default value: n/a
Source database: DOCDB
Source field name
1) Source for the standard applications:

```
<application-reference data-format="docdb" is-representative="N">
      <document-id>
                  <country>DE</country>
                  <doc-number>10331291</doc-number>
                  <kind>A</kind>
                  <date>20030710</date>
```

2) For priorities in DOCDB for which there is no application registered in DOCDB, use the authority (country) of the priority:

```
<priority-claims>
      <priority-claim sequence="1" data-format="docdb">
            <document-id>
                  <country>AE</country>
                  <doc-number>4000</doc-number>
                  <kind>A</kind>
                  <date>20000529</date>
            </document-id>
            <priority-active-indicator>Y</priority-active-indicator>
```

3) For artificial applications which were created for all artificial publications which were themselves artificially created for those cited publications, where the cited publications are not registered in DOCDB as publications: use the kind code "D2":

```
<references-cited>
      <citation srep-phase="SEA" sequence="1">
            <patcit>
                  <document-id>
                        <country>AM</country>
                        <doc-number>199</doc-number>
                        <kind>A</kind>
```

4) For artificial applications which were created for applications originating from cited applications not recorded in DOCDB: use the application kind as cited in the cited application; if not given then use "D3". Note that in 2014 Autumn Edition there was no occurrence of 'D3':

```
<citation srep-phase="APP" sequence="1">
      <patcit num="2" dnum="US19420613452A" dnum-type="application number">
            <document-id>
                  <country>US</country>
                  <doc-number>19420613452</doc-number>
                  <kind>A</kind>
                  <date>00000000</date>
            </document-id>
      </patcit>
</citation>
```

Source sub-field identifier
data-format="docdb"
Source sub-field identifier
n/a
Comments
Warning: Please consider that the application kind code landscape can be at times complicated (see also table description TLS201_APPLN in this document).
PATSTAT users must consult the DOCDB application & priority concordance documents, and the DOCDB User Documentation to avoid misinterpretation of the data.

Utility models for France have the value A, not U. To identify utility models for France, use the attribute IPR_TYPE.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 01-12-2004 - First version
R. Heijna - 21-04-2005 - Domain redefined, source extended
J. Rollinson - 18-04-2006 - Source codes extended
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 13-04-2012 - Added warning on usage of kind codes
M. Kracker - 15-03-2013 - Added artificial cited applications
M. Kracker - 01-11-2013 - Hint added to comment
M. Kracker - 01-04-2014 – Application kind P (provisional application) added
M. Kracker - 01-10-2016 – Application kind V (plant patent) added
M. Kracker - 01-10-2017 – Comment: exception for FR added
M. Kracker - 01-04-2019 – Domain description extended to cover bad data

6.11 APPLN_NR

Name: Application number
Also Known As: "Dossier number" in case of EP applications
Description: Number issued by the Patent Authority where the National, International or Regional application was filed
Domain: Up to 15 ASCII characters
This attribute must be unique in combination with APPLN_AUTH & APPLN_KIND.
The last character is either numeric or A, D, K, T or X. The DOCDB administrators make the application numbers end with a D, T or X to create "dummy" application numbers that are present because the number is mandatory but the actual number is not known.
A - data errors
D - dummy application; the publication number is put in front of the D
K – special type of older Brazilian application (number format 11nnnnnK)
T - dummy technical priority
X - dummy pre-1970 derived priority
Default value: empty string
Source database: DOCDB
Source field name
1) Source for the standard applications:

```
<application-reference data-format="docdb" is-representative="N">
      <document-id>
            <country>DE</country>
            <doc-number>10331291</doc-number>
            <kind>A</kind>
            <date>20030710</date>
```

2) Source for the artificial applications from priorities:
We assume that all priorities are accurately recorded in DOCDB. If a priority reference does not appear as an application reference, then in PATSTAT we create an artificial application with the authority (country), number, kind and date of the priority. See APPLN_ID for the rules for creating the APPLN_ID for these artificial applications.
See rules for processing PRIOR_APPLN_SEQ_NR.

```
<priority-claims>
      <priority-claim sequence="1" data-format="docdb">
            <document-id>
                  <country>AE</country>
                  <doc-number>4000</doc-number>
                  <kind>A</kind>
                  <date>20000529</date>
            </document-id>
            <priority-active-indicator>Y</priority-active-indicator>
```

3) Source for the artificial applications from citations:

```
<references-cited>
      <citation srep-phase="SEA" sequence="1">
            <patcit>
                  <document-id>
                        <country>AM</country>
                        <doc-number>199</doc-number>
                        <kind>A</kind>
```

If a cited document does not appear as a publication-reference in DOCDB, then in PATSTAT we create an artificial publication. See rules in element PUBLN_NR. We also create an artificial application, using the same country and number as the artificial publication, but we give an APPLN_FILING_DATE of 9999-12-31 and an APPLN_KIND of 'D2'. See also the rules for allocating the PUBLN_ID range.
See rules for processing CITED_PAT_PUBLN_ID.

4) For artificial applications which were created for applications originating from cited applications not recorded in DOCDB: use the application number as cited in the cited application:

```
<citation srep-phase="APP" sequence="1">
      <patcit num="2" dnum="US19420613452A" dnum-type="application number">
            <document-id>
                  <country>US</country>
                  <doc-number>19420613452</doc-number>
                  <kind>A</kind>
                  <date>00000000</date>
            </document-id>
      </patcit>
</citation>
```

Source sub-field identifier
data-format="docdb"
Source codes
n/a

Comments
The terms "Application number" and "Dossier number" are in use for the complete identification, for example "EP99101234"
See "Application Replenishment"

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 03-11-2004 - First version
R. Heijna - 20-04-2005 - Domain identified, source extended
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 15-03-2013 - Added artificial cited applications

6.12 APPLN_NR_EPODOC (deprecated)

Name: Application number in EPODOC format
Also Known As: EPODOC application number
Description: Number in EPODOC format (containing letters and digits) which, if present – will uniquely identify an application (with exceptions). The number is created by the EPO based on the DOCDB application number, application authority and application kind.
Domain: Up to 20 ASCII characters (typically, 13 - 14 characters)

Explanation of the format, according to Annex XI of the "Exchange Format" document of DOCDB, version 2.4.3 from 01.01.2013
Basic structure of application and priority-numbers in data-format="epodoc" is:

- country
- number
  - ccyy - century/year derived from application- or priority-date
  - nnnnnnn - serial number, leading zeroes when required
- kind-code, when kind-code not = 'A'

Extended structure for a number of countries:

- country ["WO" when kind-code in data-format="docdb" is "W"]
- number
  - ccyy : century/year derived from application- or priority-date
  - xx : "other data"
  - nnnnn : serial number, leading zeroes when required
- kind-code, when kind-code not = 'A'

"Other data" may be:

- regional office, e.g. 'MI' when country = 'IT' and regional office = Milan
- filing country, e.g. 'US' when country = 'WO' and filing country = US
- ...

Length of the concatenated string is generally fixed at 13 characters or 14 when the kind-code is appended. Strings exceeding a total of 13 or 14 may occur, when the number of significant digits exceeds the number of digits reserved for the serial number, e.g. DE.

A special format applies to numbers that in data-format="docdb" have been suffixed with letters 'D' or 'T' or 'X':

- country
- 'D' or 'T' or 'X'
- number
- kind-code, when kind-code not = 'A'

Default value: empty (if not provided by DOCDB due to formatting issues)
Source database: DOCDB
Source field name
1) Source for the standard (= non-artificial) applications:

```
</application-reference>
      <application-reference data-format="epodoc">
            <document-id>
                  <doc-number>US20070859929</doc-number>
            </document-id>
      </application-reference>
```

2) For all artificial applications the attribute APPLN_NR_EPODOC will contain an empty string.

Source sub-field identifier
data-format="epodoc"
Source codes
n/a

Comments
This attribute is deprecated, because the coming Espacenet version will not make use of it, but instead will use the DOCDB application number format, which is available in the attribute APPLN_NR. So the attribute APPLN_NR_EPODOC will be removed in one of the next PATSTAT editions.

The number in APPLN_NR_EPODOC is almost unique. For technical reasons there currently (April 2021) are a few thousand applications with non-unique values in APPLN_NR_EPODOC.

This attribute is useful to easily look up details on an application in Espacenet, which also uses the EPODOC application number to identify an application. You can either

- enter the attribute (e. g. DE20051040258) into the search mask of Espacenet or
- construct a URL like e. g. http://worldwide.espacenet.com/DE20051040258 to directly see the search result

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2019 – This attribute is now deprecated.

6.13 APPLN_NR_ORIGINAL

Name: Application number in original format
Also Known As: Original application number
Description: Application number in original format as provided by the supplier. It is assumed that the number is as printed on the respective publications.
Typically these numbers do not contain the authority (country) code. In about 10% of the applications no original application number is known.
Domain: Up to 100 characters
Default value: empty
Source database: DOCDB
Source field name

1) Source for the standard (= non-artificial) applications:

```
<exch:application-reference data-format="original">
      <document-id>
            <doc-number>11137814</doc-number>
      </document-id>
</exch:application-reference>
```

If DOCDB does not provide an original application number in any of the publications of an application, then APPLN_NR_ORIGINAL will contain an empty string.

If DOCDB provides multiple conflicting original application numbers for the same application, then only one (= any of the conflicting) original application numbers should be stored. (Note: This is supposed to not happen, but may still occur due to data errors)

EP publications published after 2013-03-13, the application number is published in DOCDB with a check digit, i.e. 04801606**.7**. For sake of consistency with previous original application numbers, the check digit is removed in PATSTAT.

2) For all artificial applications the attribute APPLN_NR_ORIGINAL will contain an empty string.

Source sub-field identifier
data-format="original"
Source codes
n/a
Comments
This attribute is useful to combine application data of PATSTAT with other databases which also contains the original application number.
The original application number is not necessarily unique within the same APPLN_AUTH and the same APPLN_KIND (e.g. for patents and utility models). For example, the offices of US, JP, FR, CH, CS, IT, SU seem to have re-used their application numbers at least in some periods of time.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2016 - First version

6.14 APPLN_TITLE

Name: Title of application
Also Known As: n/a
Description: Title of the application
Domain: Up to 3 000 characters
Default value: n/a
Source database: DOCDB
Source field name:
```
<invention-title lang="EN" data-format="docdba"> SURGICAL
CANNULA</invention-title>
```
Source sub-field identifier
```
data-format="docdba"
```
Comments
Only one of possibly multiple abstracts is stored. See description of table TLS203_APPLN_ABSTRACT for details.

The average size of titles is 53 characters; maximum size is 3000 for a Brazilian document (as of April 2013).

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 13-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML

6.15 APPLN_TITLE_LG

Name: Language of title of application
Also Known As: n/a
Description: Language of the title of the application selected for and loaded in PATSTAT
Domain: 2 ASCII characters, according to ISO language codes (ISO 639-1) plus these DOCDB-specific extensions for languages:

bs	Bosnian
hr	Croatian
me	Montenegrin

or spaces

Default value: spaces
Source database: DOCDB
Source field name

```
<invention-title lang="EN" data-format="docdb"> SURGICAL
CANNULA</invention-title>
```

Source sub-field identifier

```
data-format="docdb"
```

Comments
Use the value of this attribute for the title stored in the table TLS201_APPLN_TITLE.
Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 13-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker – 01-10-2018 – language codes “bs”, “hr” and “me” added

6.16 APPLT_SEQ_NR

Name: Sequence number of applicants
Also Known As: n/a
Description: Number indicating the place in the list of applicants in the application
Domain: Number 0 … about 250
Default value: 0
Source database: DOCDB
Source field name

1) EP Register for EP patent applications

2) OECD patents database for US data post 1976-01-01 up to and including November 15th, 2005 for Published Grants.

3) PATSTAT weekly file extracts from USPTO website for Published Grants from November 22nd, 2005 until today; Published Applications from September 29th, 2005 to today inclusive.

4) DOCDB Applicant sequence number for USPTO Published Applications from March 1st, 2001 to September 22nd, 2005 from DOCDB, data-format="docdba".

5) all other Applicant Sequence numbers come from DOCDB, data-format="docdba".

```
<applicant sequence="1" data-format="docdba">
      <applicant-name>
            <name>THE JOHNS HOPKINS UNIVERSITY MACDONALD, ALEX
BRUCE</name>
      </applicant-name>
</applicant>
<applicant sequence="2" data-format="docdba">
      <applicant-name>
            <name>AN, LING LING UNIVERSITY OF MASSACHUSETTS, A
PUBLIC INSTITUTION OF HIGHER EDUCATION OF THE COMMONWEALTH
OF MASSACHUSETTS,</name>
      </applicant-name>
</applicant>
```

Source sub-field identifier
`sequence="1" data-format="docdba"`
Comments
An entry with a value 1 to n represents an applicant; an entry with the value 0 does not represent an applicant, but another person (e.g. an inventor). It is possible that there are applications where no applicants are known.
Consequently, adding the condition "APPLT_SEQ_NR > 0" to the WHERE clause in a query retrieves only those persons from TLS207_PERS_APPLN or TLS227_PERS_PUBLN which are applicants.
Likewise, adding the condition "APPLT_SEQ_NR > 0 AND INVT_SEQ_NR > 0 " retrieves only persons which for a certain application are applicants as well as inventors.

For US data:
Documents published after 1976-01-01: The sequence number is designed to represent the sequence in which Applicants appear on the documents. In this database, this is accurate

for the first-named applicant. For the second- or later- named applicants, the sequence number in this database has been arbitrarily given.

For all US documents published before 1976-01-01, where the data was taken from DOCDB, the sequence numbers are believed to be correct.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 21-12-2004 - First version
R. Heijna - 07-07-2005 - Value zero for the physical model
J. Rollinson - 18-04-2006 - US data comment added
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-10-2013 - Changed source from EPO Bulletin to EP Register; changed domain
M. Kracker - 01-10-2015 - Changed comment

6.17 CITED_APPLN_ID

Name: Identification of cited application
Also Known As: n/a
Description: Surrogate key of the application that is cited
Domain: Number 0 … 999 999 999; see also attribute APPLN_ID
Default value: 0
Source database: DOCDB, PATSTAT
Source field name

Direct patent application citations:
If `citation srep-phase="APP`", indicating this citation was done by the applicant, then `citation/patcit` may contain either a reference to a cited *publication* or a reference to a cited *application*.

If `<patcit dnum-type="application number">`, then use `country`, `doc-number` and `kind` in `references-cited/citation/patcit/document-id` to find the corresponding APPLN_ID for this application via APPLN_AUTH, APPLN_NR and APPLN_KIND. The value of APPLN_ID for this application is the CITED_APPLN_ID.
APPLN_FILING_DATE is taken from the date in `citation/patcit/document-id/date`. If the date is not given then 9999-12-31 is to be used.

If there is no corresponding application in table TLS201_APPLN in PATSTAT, then create an artificial application in table TLS201_APPLN. See section 4.4.2 "Application replenishment for citations".

Usage Example EP 2305027 A2:

```
<citation srep-phase="APP" sequence="46">
      <patcit num="1" dnum="US46600890A" dnum-type="application number">
            <document-id>
                  <country>US</country>
                  <doc-number>46600890</doc-number>
                  <kind>A</kind>
                  <date>19900112</date>
            </document-id>
      </patcit>
```

Source sub-field identifier: n/a

Comments
Not only *applications* can be cited, but – much more typically - *publications* as well.

Note: Cited *publications* (see CITED_PAT_PUBLN_ID) are *not* related to cited *applications* (see CITED_APPLN_ID).

In the 2017 Autumn Edition no replenished applications having kind code "D3" occurred (see section 4.4.3 "Allocating the APPLN_ID").

Modification history
Author of update - Date of update - Explanation of update
D. Lingua - 04-08-2011 - First version

D. Lingua - 26-04-2012 - Comment on "D3" kind code added
M. Kracker - 01-04-2015 – Clarification added in comment
M. Kracker - 01-12-2015 – Patent applications can also be cited from within NPL citations
M. Kracker - 01-10-2017 – Undo the change from 01-12-2015: Patent *applications* cannot be cited from within NPL citations
M. Kracker - 01-10-2019 – Amended explanation of the source

6.18 CITED_DOCDB_FAMILY_ID

Name: ID of the cited DOCDB simple family
Also Known As: n/a
Description: Uniquely identifies the cited family. The ID has no business meaning.
Domain: Number 1 … 999 999 999
Default value: n/a
Source database: PATSTAT

Source field name: Derived from the publication information (TLS211_PAT_PUBLN), citation information (TLS212_CITATION) and DOCDB family information (TLS201_APPLN)

Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version

6.19 CITED_NPL_PUBLN_ID

Name: Identification of cited non-patent literature
Also Known As: n/a
Description: Surrogate key for Non-Patent Literature publications which has been cited
Domain: 32 ASCII characters, or the digit 0

Default value: 0
Source database: DOCDB, PATSTAT
Source field name:
See attribute NPL_PUBLN_ID

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2016 – Name changed (was: NPL_PUBLN_ID)
M. Kracker - 01-04-2021 – Domain has changed from numeric to string

6.20 CITED_PAT_PUBLN_ID

Name: Identification of cited patent publication
Also Known As: n/a
Description: Surrogate key of the publication that is cited
Domain: Number 0 … 999 999 999; see also attribute PAT_PUBLN_ID
Default value: 0
Source database: DOCDB, PATSTAT
Source field name

1) Direct patent publication citations:
If `<patcit dnum-type="publication number>`, then use `country`, `doc-number` and `kind` in `references-cited/citation/patcit/document-id` in DOCDB to find the corresponding PAT_PUBLN_ID for this publication in in PATSTAT via PUBLN_AUTH, PUBLN_NR and PUBLN_KIND. The value of PAT_PUBLN_ID for this publication is the CITED_PUBLN_ID.

If there is no corresponding publication in PAT_PUBLN in PATSTAT, an artificial publication in table TLS211_PAT_PUBLN is to be created. Besides the key-elements, PUBLN_DATE is filled from `citation/patcit/document-id/date`, if it is present.
An artificial application must then be created as well, with APPLN_AUTH equal PUBLN_AUTH, APPLN_NR equal PUBLN_NR and APPLN_KIND equal 'D2'. The APPLN_FILING_DATE is the same as the PUBLN_DATE of the corresponding artificial publication.
A corresponding surrogate key APPLN_ID must also be created, in the range of IDs for artificial applications for artificial cited publications.

2) Patent publications cited within Non-Patent Literature citations:
Here CITED_PAT_PUBLN_ID refers to a patent publication id which has been extracted from a Non-Patent Literature citation. In a row in table TLS212_CITATION, you will find these columns populated:

- PAT_PUBLN_ID
- CITN_ID
- CITN_ORIGIN
- CITED_PAT_PUBLN_ID
- CITED_NPL_PUBLN_ID
- NPL_CITN_SEQ_NR

If the CITED_NPL_PUBLN_ID is not 0, and if that NPL citation refers to a patent document, then CITED_PAT_PUBLN_ID will hold the value of the PAT_PUBLN_ID of the referenced patent document.
If the referenced patent document cannot be found as a publication-reference in DOCDB, then create an artificial publication for it (see case 1) above).

The referenced patent document is the document defined in the element `references-cited/citation/nplcit/source-doc/document-id` in DOCDB . There is at most one `<source-doc>` element.

The `<source-doc>` element will always contain one reference to a patent *publication*, and never a reference to a patent *application*.

Usage Example:

```
<references-cited>
      ...
      <citation cited-phase="SEA" cited-date="20110103" srep-office="EP"
            sequence="4">
            <nplcit num="1" npl-type="d" extracted-xp="002391653">
                  <text>DATABASE WPI Week 200577, Derwent Publications Ltd.,
London, GB; AN 2005-752331, XP002391653</text>
                  <online>
                        <edition>0</edition>
                        <vid>2005</vid>
                        <ino>77</ino>
                        <absno>2005-752331</absno>
                  </online>
                  <source-doc>
                        <document-id>
                              <country>JP</country>
                              <doc-number>2005281133</doc-number>
                              <kind>A</kind>
                        </document-id>
                  </source-doc>
            </nplcit>
      </citation>
</references-cited>
```

Source sub-field identifier
n/a
Comments
No self-citing is allowed, so ignore any cited documents which are the same as the publication-reference. In this respect, ignore the Kind Code for EP publications. I.e. if EP1000000B1 cites EP1000000 with any kind code (including EP1000000 with no kind code), then ignore this citation.

An aggregate count of publications in PATSTAT will result in a higher count than in DOCDB, due to the inclusion of these artificial publications in PATSTAT. The difference is usually at the publication kind code level, as the cited kind code is incomplete or missing. For example, publication EP1000000A in PATSTAT is artificial, it does not exist in DOCDB - the correct kind code is A1, e.g. EP1000000A1

See also the rules in the description of table TLS212_CITATION.

Note: Cited publications (see CITED_PAT_PUBLN_ID) are *not* related to cited applications (see CITED_APPLN_ID).

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 04-05-2005 - First version
R. Heijna - 15-11-2005 - Special EP rule removed
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-04-2015 – Clarification added in comment
M. Kracker - 01-12-2015 – New processing rules for citations within NPL citations (case 2)

M. Kracker - 01-04-2018 – Simplification of section “Source field name”.
“Corresponding docs” are not considered anymore.
M. Kracker - 01-10-2019 – Amended explanation of the source (case 2)

6.21 CITN_CATEG

Name: Categories of the citation
Also Known As: n/a
Description: Categories of the citation as mentioned in Search Reports
Domain: Up to 10 ASCII characters (typically X, I, Y, A, D, E, P, L, R, T, O, &)
See DOCDB manual Annex XIV (https://www.epo.org/searching-for-patents/data/bulk-data-sets/manuals.html)
Default value: n/a
Source database: DOCDB
Source field name

```
<exch:citation cited-phase="SEA" cited-date="20120112" srep-office="EP" sequence="1">
	<patcit num="1" dnum="WO2010141409A1" dnum-type="publication number">
		<document-id doc-id="329547194">
			<country>WO</country>
			<doc-number>2010141409</doc-number>
			<kind>A1</kind>
			<name>CRONIN MICHAEL D [US], et al</name>
			<date>20101209</date>
		</document-id>
	</patcit>
	<rel-passage>
		<passage>
			<para>108</para>
			<figure>23A,23B</figure>
		</passage>
		<category>XP</category>
		<rel-claims>1-4,6,8,9,11</rel-claims>
		<category>I</category>
		<rel-claims>5</rel-claims>
	</rel-passage>
	<category>XPI</category>
</exch:citation>
```

For “poor” citations the data is taken from the XML element
`citation\category`.
For “rich” citations the data is taken from the XML element
`citation\rel-passage\category`.

Source sub-field identifier
n/a

Comments
For “poor” citations (i.e. citations without rich structure) CITN_CATEG contains all citation categories of this citation as a single string, like “YXPI”, regardless to which claims of the examined applications they refer.
In “rich” citations the citation categories always refer to a specific set of claims. So citation categories which are not applied to the same claim will not be in the same CITN_CATEG value. So for example, “Y” might refer to the claims 1-3 whereas “XP” might refer to the claims 6 and “I” might refer to claim 14.

Only when CITN_ORIGIN is SEA, ISR, SUPP and PRS (= citations introduced during search, International Search Report, Supplementary Search Report or pre-search) categories may – but need not - occur; in general only the search examiners give these categories. For some countries (e. g. US, JP, but also other countries) no categories are available.

See Annex XIV of the DOCDB User Documentation for an explanation of the meaning of the categories.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 06-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 11-10-2010 - Added rules on SEA categories
D. Lingua - 26-04-2012 - Added category "I"
M. Kracker – 01-04-2014 – Clarified comment
M. Kracker – 01-04-2015 – Comment changes: Restrictions on number of categories per application has been lifted
M. Kracker – 01-04-2015 – Comment changes: Categories may also occur in the PRS phase.
M. Kracker – 01-04-2018 – Citation category “&” has been added
M. Kracker – 01-04-2019 – General changes; CITN_CATEG can now hold multiple categories

6.22 CITN_GENER_AUTH

Name: Identification of the International Search Authority (ISA) for PCT search reports (incl. supplementary search reports) or the national/regional search authority in other cases
Also Known As: n/a
Description: Country code of the (Supplementary) International Search Authority (ISA / SISA) for PCT search reports (incl. supplementary search reports) or the national/regional search authority in other cases
Domain: 2 characters (A-Z), according to WIPO ST.3 or spaces
Default value: spaces
Source database: DOCDB
Source field name `< citation srep-phase="ISR" srep-office="AT" sequence="1">`

Usage Example:

```
<references-cited>
      <citation srep-phase="ISR" srep-office="AT" sequence="1">
            <patcit num="1" dnum="US4996335A" dnum-type="publication number">
                  <document-id>
                        <country>US</country>
                        <doc-number>4996335</doc-number>
                        <kind>A</kind>
                        <date>19910226</date>
                  </document-id>
            </patcit>
            <category>X</category>
      </citation>
      <citation srep-phase="ISR" srep-office="AT" sequence="2">
            <patcit num="2" dnum="BE889563A1" dnum-type="publication number">
                  <document-id>
                        <country>BE</country>
                        <doc-number>889563</doc-number>
                        <kind>A1</kind>
                        <date>19811103</date>
                  </document-id>
            </patcit>
            <category>X</category>
      </citation>

      ...
</references-cited>
```

Source sub-field identifier: n/a

Comments
The column CITN_GENER_AUTH will only be populated where CITN_ORIGIN is

- ISR or SUP, then identifying the (Supplementary) International Search Authority (ISA)

  or
- SEA, EXA or PRS,
  then identifying a national / regional search authority.

Modification history
Author of update **-** Date of update - Explanation of update

D. Lingua - 04-08-2011 - First version
M. Kracker - 01-10-2016 – Comment changed
M. Kracker - 01-04-2017 – Description and comment changed
M. Kracker - 01-10-2018 – Comment amended

6.23 CITN_ID

Name: Citation identification
Also Known As: n/a
Description: Number distinguishing the citations in one citing document (patent publication)
Domain: Number 1 .. about 1100
Default value: n/a
Source database: Computed from PATSTAT. It is a sequential number for each citation within one citing patent publication. The numbering starts with 1.

Source field name: n/a
Source sub-field identifier: n/a
Comments
The number does not bear a particular meaning; it is just a running number among all citations in one citing document.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 21-11-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-12-2015 – Domain and processing instructions changed

6.24 CITN_ORIGIN

Name: Origin of the citation
Also Known As: Citation phase
Description: Provenance of the citation
Domain: 3 ASCII character code

The code indicates the origin of the citation:

APP citations introduced by the applicant

SEA citations introduced during search (from Search Report)

ISR citations from the International Search Report

SUP citations from the Supplementary Search Report

PRS "PRe-Search" citations (available before official publication; only for US applications; further details see *Comments* section below)

EXA citations introduced during examination

OPP the real opposition documents (citations) selected by the opposition division (published with a European Patent Specification (EP-B2))

APL citations introduced when filed for appeal by applicant / proprietor / patentee

FOP when an opposition has been filed: citations introduced by the opponent or the proprietor.

TPO citations introduced because of Third Party Observations (Art 115 EPC)

CH2 citations introduced during the Chapter 2 phase of the PCT

Default value: n/a

Source database: DOCDB
Source field name

```
<citation srep-phase="SEA" sequence="1">
      <patcit num="1" dnum="WO9505670A1" dnum-type="publication number">
            <document-id>
                  <country>WO</country>
                  <doc-number>9504670</doc-number>
                  <kind>A1</kind>
            </document-id>
      </patcit>
      <category>Y</category>
</citation>
<citation srep-phase="SEA" sequence="2">
      <patcit num="2" dnum="DE4135041A1" dnum-type="publication number">
            <document-id>
                  <country>DE</country>
                  <doc-number>4135041</doc-number>
                  <kind>A1</kind>
            </document-id>
      </patcit>
      <category>Y</category>
</citation>
<citation srep-phase="SEA" sequence="3">
      <patcit num="3" dnum="FR2730035A1" dnum-type="publication number">
            <document-id>
                  <country>FR</country>
                  <doc-number>2730035</doc-number>
                  <kind>A1</kind>
            </document-id>
```

```
      </patcit>
      <category>Y</category>
</citation>
<citation srep-phase="APP" sequence="1">
      <patcit num="1" dnum="DE4007646A1" dnum-type="publication number">
            <document-id>
                  <country>DE</country>
                  <doc-number>4007646</doc-number>
                  <kind>A1</kind>
            </document-id>
      </patcit>
</citation>
<citation srep-phase="APP" sequence="2">
      <patcit num="2" dnum="JP4241100A" dnum-type="publication number">
            <document-id>
                  <country>JP</country>
                  <doc-number>4241100</doc-number>
                  <kind>A</kind>
            </document-id>
      </patcit>
</citation>
<citation srep-phase="APP" sequence="3">
      <patcit num="3" dnum="JP7044800A" dnum-type="publication number">
            <document-id>
                  <country>JP</country>
                  <doc-number>7044800</doc-number>
                  <kind>A</kind>
            </document-id>
      </patcit>
</citation>
```

Comments
Look at the value of the element srep-phase to get the value of CITN_ORIGIN.
Element `<patcit>` may contain cited publications or cited applications. Cited applications only when `srep-phase="APP"`.

The table “Overview of citation data in the EPO's citation database (REFI)” in https://www.epo.org/searching-for-patents/data/coverage/regular.html provides a full list of origins available for a given authority.

The term *Pre-Search* is used by the EPO for a sort of search done by the USPTO examiner during the examination procedure which is independent of the publication rhythm. Due to bilateral agreements the EPO receives these search results from the USPTO.
Once the US “A” publication appears, the EPO appends these citations to this “A” publication.
In some cases these searches are done later than the A publications, so the citations could cite documents which are published *after* the publication date of the A publication.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 06-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 14-06-2010 - Introduced new citation origins
D. Lingua - 08-10-2012 - Introduced new citation origin PRS
M. Kracker - 01.10.2013 - Changed domain; Introduced new citation origins;

clarification of codes
M. Kracker - 01.10.2015 - Code 115 changed to code TPO
M. Kracker - 01.04.2019 – Comment amended

6.25 CITN_REPLENISHED

Name: Replenished citation indicator
Also Known As: n/a
Description: the PAT_PUBLN_ID of the international publication from which the citation has been copied to an EP publication; 0 otherwise (no replenishment)
Domain: number 0 ... 999 999 999
Default value: 0
Source database: PATSTAT
Source field name: Derived from table TLS212_CITATION, from attributes APPLN_AUTH and INTERNAT_APPLN_ID of table TLS201_APPLN and from table TLS216_APPLN_CONT.

Source sub-field identifier: n/a
Comments:
A search report for a Euro-PCT application will in most cases *not* repeat the citations which are already in the international search report for the PCT application (cf. “Guidelines for Examination in the EPO”, section X-9.1.4). Consequently, when analysing citations of Euro-PCT applications, one would also need to include the citation of the corresponding PCT application

To avoid this potential pitfall, in PATSTAT the citations of Euro-PCT publications are replenished with the citations of their international publication. Euro-PCT applications which are some sort of continuation (see table TLS216_APPLN_CONT) of another application are not considered for replenishment.
So the applications which are considered are identified by:

- APPLN_AUTH = EP
- INTERNAT_APPLN_ID > 0
- the APPLN_ID is not in APPLN_ID of the table TLS216_APPLN_CONTN

Of these applications, their publications are replenished like this:

- **A1** publications of the Euro-PCT application are replenished by the citations of the **A1** publications of the corresponding PCT application.
- **A2** publications of the Euro-PCT application are replenished by the citations of the **A2 and A3** publications of the corresponding PCT application.

The attribute CITN_REPLENISHED can be used to identify those replenished citations whose origin is the international publication.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2018 - First version
M. Kracker - 01-04-2019 – Comment and logic of computation amended

6.26 CITY

Name: City part of the address
Also Known As: n/a
Description: Contains the city part of the address
Domain: Up to 200 characters
Default value: empty string
Source database: USPTO data of published applications and published grants

Source field name:
`<addressbook> <address> <city>`
Source sub-field identifier: n/a

Comments:
n/a
Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-10-2015 – Removed source “EP Register data”;
cf. attributes ADDRESS_1, ..., ADDRESS_5

6.27 CLASS_SCHEME

Name: Scheme of the classification
Also Known As: n/a
Description: Scheme of the corrected classification
Domain: up to 4 ASCII characters. Possible values:
- IPC
- empty string

Default value: empty string

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/patent-classification/classification-scheme/@scheme

```
<legal-event providing-office="EP" date-added="20110505" date-previous-
exchange="20110505" sequence-number="3">
     <event-date>20110504</event-date>
     <event-code>RIC1</event-code>
     <event-details>
          <event-description event-description-type="original">KLASSIFIKATION
(KORR.)</event-description>
          <event-description lang="en">CLASSIFICATION (CORRECTION)</event-
description>
          <patent-classification>
               <classification-scheme scheme="IPC"/>
               <classification-symbol>G09G   3/32
20060101AFI20110331BHEP</classification-symbol>
          </patent-classification>
     </event-details>
</legal-event>
```

Comments
Information about corrections of classifications is very sparse.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.28 CLASS_SYMBOL

Name: Classification symbol
Also Known As: n/a
Description: Corrected classification symbol, in WIPO ST.8 format.
Domain: up to 50 ASCII characters.
The format of the IPC symbol varies. Some are structured according to WIPO ST.8, others are unstructured.
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/patent-classification/classification-symbol

```
<legal-event providing-office="EP" date-added="20110505" date-previous-
exchange="20110505" sequence-number="3">
     <event-date>20110504</event-date>
     <event-code>RIC1</event-code>
     <event-details>
          <event-description event-description-type="original">KLASSIFIKATION
(KORR.)</event-description>
          <event-description lang="en">CLASSIFICATION (CORRECTION)</event-
description>
          <patent-classification>
               <classification-scheme scheme="IPC"/>
               <classification-symbol>G09G   3/32
20060101AFI20110331BHEP</classification-symbol>
          </patent-classification>
     </event-details>
</legal-event>
```

Comments
Information about corrections of classifications is very sparse.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.29 CONTINENT

Name: Continent
Also Known As: n/a
Description: Name of the continent (in English) in which a state is located (according to Wikipedia)
Domain: Up to 25 ASCII characters:

Only populated if the attribute STATE_INDICATOR is 'Y':
May have one of these 6 values:

- Africa
- Asia
- Australia and Oceania
- Europe
- Europe/Asia
  Note that the Russian Federation, Soviet Union and Turkey have the value 'Europe/Asia'
- North America
- South America

Default value: empty
Source database: based on Wikipedia
Source field name: n/a
Source sub-field identifier: n/a
Comments:

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version
M. Kracker - 01-04-2017 – America split into North America and South America

6.30 CONTN_TYPE

Name: Continuation type
Also Known As: n/a
Description: The type of continuation describing what relation the later application has to the earlier application. In DOCDB, this is known as the type of linkage between applications and priorities.
Domain: 3 ASCII characters
ADD - Addition
CON - Continuation
CGT - Cognate
CIP - Continuation in part
DIV - Division
INN - Internal priority
P2U - Patent which has been changed into a utility model
REI - Re-issue
SBS - Substitute
SUP - Supplementary disclosure
U2P - Utility model which has been changed into a patent
spaces - unknown

Default value: n/a

Source database: DOCDB
Source field name

```
<priority-claims>
      <priority-claim sequence="1" data-format="docdb" status="A">
            <country>US</country>
            <doc-number>90976604</doc-number>
            <kind>A</kind>
            <date>20040802</date>
            <priority-active-indicator>Y</priority-activeindicator>
      </priority-claim>
      <priority-claim sequence="1" data-format="epodoc">
            <doc-number>US20040909766</doc-number>
      </priority-claim>
      <priority-claim sequence="2" data-format="docdb" status="A">
            <country>US</country>
            <doc-number>9885602</doc-number>
            <kind>A</kind>
            <date>20020314</date>
            <priority-linkage-type>3</priority-linkage-type>
            <priority-active-indicator>N</priority-activeindicator>
      </priority-claim>
      <priority-claim sequence="2" data-format="epodoc">
            <doc-number>US20020098856</doc-number>
      </priority-claim>
      <priority-claim sequence="1" data-format="original">
            <doc-number>9885602/doc-number>
      </priority-claim>
</priority-claims>
```

From the application publication authority code (APPLN_AUTH) and the `priority-linkage-type` the continuation type is determined from the table in section 4.6.2 Continuation types.

Note:
a) if there is no element `<priority-linkage-type>`, then put spaces in CONTN_TYPE.

b) if there is no matching entry in the table, then put spaces in CONTN_TYPE.

Note that before 1991, the EPO did not record the so called "linkage type" of priority numbers, that is the EPO did not record which kind of relation a given priority number has (Paris Union priority, continuation, division, etc.). Data in this element prior to 1991 is thus not reliable.

Source sub-field identifier
data-format="docdb"

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 04-05-2005 - First version
R. Heijna - 13-07-2005 - Domain adapted
J. Rollinson - 14-02-2008 - pre 1991 US fact.
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-04-2020 – Domain extended by P2U and U2P

6.31 CPC_ACTION_DATE

Name: CPC action date
Also Known As: n/a
Description: The date of assigning the classification symbol
Domain: Date between '2013-01-01' and current date
Default value: n/a
Source database: DOCDB
Source field name:

```
<patent-classifications>
  <patent-classification sequence="1">
    <classification-scheme office="EP" scheme="CPCI">
      <date>20130101</date>
    </classification-scheme>
    <classification-symbol>G01S  11/16        </classification-symbol>
    <symbol-position>F</symbol-position>
    <classification-value>I</classification-value>
    <classification-status>B</classification-status>
    <classification-data-source>H</classification-data-source>
    <generating-office>EP</generating-office>
    <action-date>
      <date>20151127</date>
    </action-date>
  </patent-classification>
```

Source sub-field identifier: n/a

Comments
See also table TLS225_DOCDB_FAM_CPC (CPC symbols assigned on DOCDB family level) and table TLS224_APPLN_CPC (CPC symbol redundantly stored on application level).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2020 - First version

6.32 CPC_CLASS_SYMBOL

Name: CPC classification symbol
Also Known As: CPC class, CPC classification, CPC symbol
Description: Classification symbol according to the Cooperative Patent Classification
Domain: Up to 19 characters (A-Z, 0-9, /, space);
All values which are allowed by the CPC;
Corresponds to position 1 - 19 (i.e. section, class, subclass, main group, subgroup) of the 50 character long text string as defined by WIPO ST.8 with trailing spaces removed.
Examples: `A61K`
`H04Q  7/32`
`C07K 14/00`
`C07D 405/06`
`H01M2220/20`
Note that spaces may be required on position 5-7, because the slash "/" is always on the 9th position. For more details see the table below.
Default value: n/a
Source database: DOCDB
Source field name:

```
<patent-classifications>
  <patent-classification sequence="1">
    <classification-scheme office="EP" scheme="CPCI">
      <date>20130101</date>
    </classification-scheme>
    <classification-symbol>G01S  11/16       </classification-symbol>
    <symbol-position>F</symbol-position>
    <classification-value>I</classification-value>
    <classification-status>B</classification-status>
    <classification-data-source>H</classification-data-source>
    <generating-office>EP</generating-office>
    <action-date>
      <date>20151127</date>
    </action-date>
  </patent-classification>
```

Source sub-field identifier
without trailing spaces

Recording of IPC (CPC is compatible to IPC) is described in WIPO ST.8:
For the recording of CPC symbols on machine-readable records a field of 50 positions should be allotted for each symbol, the 50 positions of the field to be used as follows:

symbol, the 50 positions of the field to be used as follows: *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
9	Separating character	/ (“Slash”)
10 to 15	Subgroup (left aligned)	00,…,999999, blank

16 to 19	For future use	4 blanks
20 to 27	Version indicator	YYYYMMDD date format
28	core / advanced	not applicable
29	First or later position of symbol	F,L
30	Classification value (invention or additional)	I,A
31 to 38	Action date	YYYYMMDD date format
39	Original or reclassified data	B,R
40	Source of classification data (human, machine, generated)	H, M, G
41-42	Generating office	AA,…,ZZ (ST.3)
43-50	For future use	8 blanks

For each symbol, be sure to take the corresponding values of CPC_GENER_AUTH, CPC_VERSION, CPC_POSITION and CPC_VALUE from the same `patent_classification` element.

Comments
See also table TLS225_DOCDB_FAM_CPC (CPC symbols assigned on DOCDB family level) and table TLS224_APPLN_CPC (CPC symbol redundantly stored on application level).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 13-03-2013 - First version
M. Kracker - 15-10-2014 - Comments updated
M. Kracker - 01-04-2016 – Examples showing the correct format have been added
M. Kracker - 01-04-2020 – Comment changed; source example updated

6.33 CPC_DATA_SOURCE

Name: Source of CPC classification data
Also Known As: n/a
Description: Source of CPC classification data
Domain: 1 character: H, C or G:

- H Human generated data (intellectual classification by persons)
- C Classification by concordance, e.g. by copying symbols allocated by other patent offices, or by copying IPC symbols into CPC allocations
- G Classification symbols generated by software using automatic analysis of the content of the patent document

Default value: n/a
Source database: DOCDB
Source field name:

```
<patent-classifications>
  <patent-classification sequence="1">
    <classification-scheme office="EP" scheme="CPCI">
      <date>20130101</date>
    </classification-scheme>
    <classification-symbol>G01S  11/16        </classification-symbol>
    <symbol-position>F</symbol-position>
    <classification-value>I</classification-value>
    <classification-status>B</classification-status>
    <classification-data-source>H</classification-data-source>
    <generating-office>EP</generating-office>
    <action-date>
      <date>20151127</date>
    </action-date>
  </patent-classification>
```

Source sub-field identifier: n/a

Comments
See also table TLS225_DOCDB_FAM_CPC (CPC symbols assigned on DOCDB family level) and table TLS224_APPLN_CPC (CPC symbol redundantly stored on application level).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2020 - First version

6.34 CPC_GENER_AUTH

Name: CPC generating authority
Also Known As: n/a
Description: Patent office that classified the application with a CPC symbol
Domain: up to 2 characters (A-Z) according to WIPO ST.3
Default value: n/a
Source database: DOCDB

```
<patent-classifications>
  <patent-classification sequence="1">
    <classification-scheme office="EP" scheme="CPCI">
      <date>20130101</date>
    </classification-scheme>
    <classification-symbol>G01S  11/16        </classification-symbol>
    <symbol-position>F</symbol-position>
    <classification-value>I</classification-value>
    <classification-status>B</classification-status>
    <classification-data-source>H</classification-data-source>
    <generating-office>EP</generating-office>
    <action-date>
      <date>20151127</date>
    </action-date>
  </patent-classification>
```

Source field name

```
    <generating-office>GB</generating-office>
```

Source sub-field identifier: n/a

Comments
See also table TLS225_DOCDB_FAM_CPC (CPC symbols assigned on DOCDB family level) and table TLS224_APPLN_CPC (CPC symbol redundantly stored on application level).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 13-03-2013 - First version
M. Kracker - 15-10-2014 - Comments updated
M. Kracker - 01-10-2019 - CPC_GENER_AUTH is also populated for CPC_SCHEME=”CPC”
M. Kracker - 01-04-2020 – Comment changed; source example updated; Modification from 01-10-2019 is not relevant anymore

6.35 CPC_POSITION

Name: First or later position of CPC symbol
Also Known As: n/a
Description: Indicates the position of the class symbol in the sequence of classes that form the classification.
Domain: 1 character; F = first, L = later, space = unidentified
Default value: space
Source database: DOCDB

```
<patent-classifications>
  <patent-classification sequence="1">
    <classification-scheme office="EP" scheme="CPCI">
      <date>20130101</date>
    </classification-scheme>
    <classification-symbol>G01S  11/16        </classification-symbol>
    <symbol-position>F</symbol-position>
    <classification-value>I</classification-value>
    <classification-status>B</classification-status>
    <classification-data-source>H</classification-data-source>
    <generating-office>EP</generating-office>
    <action-date>
      <date>20151127</date>
    </action-date>
  </patent-classification>
```

Source field name:
<symbol_position>L</symbol_position>
This field is only available for scheme "CPC".
This field is not used with scheme "CPCNO".

Source sub-field identifier: n/a

Comments
The following facts are asserted by DOCDB:

- only one CPC allocated by a given patent office to a given patent family will be identified to have symbol-position = "F" (first)
- CPC symbol identified by symbol-position = "F" (first) will always have classification-value = "I" (invention)
- the most recent CPC symbol allocated by the USPTO will be identified "first"
- failing the presence of a USPTO allocated CPC that can be identified "first", the most recent CPC symbol allocated by the EPO will be identified "first"
- all other CPC symbols allocated to a given patent family - whether USPTO or EPO, whether invention or additional - will have symbol-position = "L" (later)

For patent authorities where the law entails the concept of "first" class, the first class symbol in a list of class symbols is the main class. For other authorities, like the EPO, there is no meaning in the position - classes may be quoted in alphabetical order for instance. Some researchers use a weighting technique to analyse by CPC.

See also table TLS225_DOCDB_FAM_CPC (CPC symbols assigned on DOCDB family level) and table TLS224_APPLN_CPC (CPC symbol redundantly stored on application level).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 13-03-2013 - First version
M. Kracker - 15-10-2015 – Comment updated
M. Kracker - 01-04-2020 - Comment changed; source example updated
V. Hassler - 28.09.2021 - Description and first comment corrected

6.36 CPC_STATUS

Name: Original or reclassified CPC data
Also Known As: n/a
Description: Indication whether the CPC is as originally assigned or whether and how it has been reclassified
Domain: 1 character: B or R:

- B basic or original data
  Original data is the first data assigned to the document.
- R reclassified data
  Reclassified data is data changed due to a change in the classification schemes.

Default value: n/a
Source database: DOCDB
Source field name:

```
<patent-classifications>
  <patent-classification sequence="1">
    <classification-scheme office="EP" scheme="CPCI">
      <date>20130101</date>
    </classification-scheme>
    <classification-symbol>G01S  11/16        </classification-symbol>
    <symbol-position>F</symbol-position>
    <classification-value>I</classification-value>
    <classification-status>B</classification-status>
    <classification-data-source>H</classification-data-source>
    <generating-office>EP</generating-office>
    <action-date>
      <date>20151127</date>
    </action-date>
  </patent-classification>
```

Source sub-field identifier: n/a

Comments
See also table TLS225_DOCDB_FAM_CPC (CPC symbols assigned on DOCDB family level) and table TLS224_APPLN_CPC (CPC symbol redundantly stored on application level).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2020 - First version

6.37 CPC_VALUE

Name: Classification value
Also Known As: Invention / Additional
Description: Indication of the value of the classification i.e. is the class symbol relating to the invention or to aspects not related to the invention (but in the application).
Domain: 1 character; I=Invention A=Additional (Non-invention)
Default value: n/a
Source database: DOCDB
Source field name:

```
<patent-classifications>
  <patent-classification sequence="1">
    <classification-scheme office="EP" scheme="CPCI">
      <date>20130101</date>
    </classification-scheme>
    <classification-symbol>G01S  11/16        </classification-symbol>
    <symbol-position>F</symbol-position>
    <classification-value>I</classification-value>
    <classification-status>B</classification-status>
    <classification-data-source>H</classification-data-source>
    <generating-office>EP</generating-office>
    <action-date>
      <date>20151127</date>
    </action-date>
  </patent-classification>
```

Source sub-field identifier: n/a

Comments
See also table TLS225_DOCDB_FAM_CPC (CPC symbols assigned on DOCDB family level) and table TLS224_APPLN_CPC (CPC symbol redundantly stored on application level).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 13-03-2013 - First version
M. Kracker - 15-10-2015 – Comment updated
M. Kracker - 01-04-2020 – Comment changed; source example updated

6.38 CPC_VERSION

Name: CPC version
Also Known As: n/a
Description: Version of the CPC
Domain: Date between '2013-01-01' and current date
Default value: n/a
Source database: DOCDB
Source field name:

```
<patent-classifications>
  <patent-classification sequence="1">
    <classification-scheme office="EP" scheme="CPCI">
      <date>20130101</date>
    </classification-scheme>
    <classification-symbol>G01S  11/16        </classification-symbol>
    <symbol-position>F</symbol-position>
    <classification-value>I</classification-value>
    <classification-status>B</classification-status>
    <classification-data-source>H</classification-data-source>
    <generating-office>EP</generating-office>
    <action-date>
      <date>20151127</date>
    </action-date>
  </patent-classification>
```

Source sub-field identifier: n/a

Comments
See also table TLS225_DOCDB_FAM_CPC (CPC symbols assigned on DOCDB family level) and table TLS224_APPLN_CPC (CPC symbol redundantly stored on application level).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 13-03-2013 - First version
M. Kracker - 15-10-2014 – Comment updated
M. Kracker - 01-04-2020 – Comment changed; source example updated

6.39 CTRY_CODE

Name: Country code
Also Known As: cc, country, territory, state, office, authority
Description: The two-letter code for the representation of states, other entities and intergovernmental organisations, as defined in WIPO standard ST.3 (plus minor additions)
Domain: 2 ASCII characters
Default value: n/a
Source database: WIPO standard ST.3
Source field name: n/a
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version

6.40 DESIGNATED_STATES

Name: Designated state(s)
Also Known As: n/a
Description: List of two-letter codes of designated states
Domain: up to 1 000 ASCII characters, consisting of an alphabetically ordered list of 2-character codes (according to WIPO ST.3), with each code separated by a comma ",":
Examples: "FR" or "AT,DE" or " DE,FR,GB,NL"
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/designated-states/country

```
<legal-event providing-office="EP" date-added="20111103" date-previous-
exchange="20111103" sequence-number="10">
     <event-date>20111102</event-date>
     <event-code>AK</event-code>
     <event-details>
          <event-description event-description-type="original">BENANNTE
VERTRAGSSTAATEN</event-description>
          <event-description lang="en">DESIGNATED CONTRACTING STATES:</event-
description>
          <event-reference>
               <event-ref-kind>
                    <kind>B1</kind>
               </event-ref-kind>
          </event-reference>
          <designated-states>
               <country>AT</country>
               <country>BE</country>
...
               <country>SM</country>
               <country>TR</country>
          </designated-states>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.41 DISCONTINUED

Name: Indicator whether a state or organisation no longer exists.
Also Known As: n/a
Description: Indicator whether a state or organisation no longer exists (according to WIPO standard ST.3).
Domain: 1 ASCII character: Y or space;

Y	discontinued
space	otherwise

Default value: n/a
Source database: WIPO ST.3
Source field name: n/a
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version

6.42 DOCDB_FAMILY_ID

Name: Identifier of a DOCDB simple family
Also Known As: DOCDB family ID; Simple family ID
Description:
A DOCDB family means that most probably the applications share exactly the same priorities (Paris Convention or technical relation or others) as contained in table TLS201_APPLN, TLS204_PRIOR_APPLN, TLS205_TECH_REL and TLS216_APPLN_CONTN.
Domain: Number 0 … 999 999 999
Default value: n/a
Source database: DOCDB
Source field name

```
<exchange-document country="DE" doc-number="10331291" kind="A1" family-id="
33441709" date="20050217" is-representative="Y" date-of-last-
exchange="2006120611" date-of-previous-exchange="20050217" date-added-
docdb="20050201" status="A">
```

In addition: For the dummy application (i.e. APPLN_ID = 0) and for artificial applications (i.e. APPLN_ID > 900 000 000) the value of the DOCDB_FAMILY_ID will be the same as the value of the APPLN_ID.

Source sub-field identifier

```
family-id
```

Comments

Every application belongs to exactly one DOCDB family. In the trivial case, an application belongs to a DOCDB family which consists of no other family members except this application itself. This is e.g. the case for all artificial applications (APPLN_ID > 900 000 000; see section 4.4).

Generally speaking, if two applications claim exactly the same prior applications as priorities (these can be e. g. Paris Convention priorities or technical relation priorities – for details see section 4.4.1 “Application replenishment for priorities”), then they are defined by the EPO as belonging to the same DOCDB simple family. The EPO reserves the right to classify an application into a particular simple family irrespective of this general rule - the EPO does this by creating artificial priorities for an application or by ignoring certain priorities (declaring them “inactive”) for the purpose of family building.

The simplified definition of the DOCDB family is that all their priorities must be the same. DOCDB family members generally refer to the same invention.

The simple family is also at times used to attribute automatically the same CPC classification symbols and other attributes to their family members.

As a general rule, the value of the DOCDB_FAMILY_ID will not change. It will be the same across editions of DOCDB and PATSTAT. However, corrections to priority numbers or changes in the priority pictures (priority numbers changing from active to inactive or vice-versa) might lead to a change in the family-ID of a given publication. See also section 4.3.2 “Stable IDs”.

Modification history
Author of update - Date of update - Explanation of update
J. Rollinson - 13-03-2008 - First version
D. Lingua - 14-05-2008 - Revised text
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 02-10-2013 - Extended for use in table TLS201_APPLN (PATSTAT Online Extension)
M. Kracker - 01-04-2015 – Revised comments; attribute in PATSTAT Online deprecated
M. Kracker - 01-10-2015 – Revised comments
M. Kracker - 01-10-2016 – Revised comments

6.43 DOCDB_FAMILY_SIZE

Name: Size of DOCDB simple family
Also Known As: n/a
Description: Size of DOCDB simple family of a given application
Domain: Number 1 … about 1.000
Default value: n/a
Source database: PATSTAT
Source field name: Derived from table TLS201_APPLN

Source sub-field identifier: n/a

Comments:
A family size of 1 means that the application is the only member in that family.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-10-2015 – Source changed to TLS201_APPLN
M. Kracker - 01-10-2016 – Change of Domain. Minimum value is 1 (was: 0)

6.44 DOC_STD_NAME

Name: Standardised name as recorded in DOCDB
Also Known As: n/a
Description: Standard name attributed to applicant and inventor names for inclusion in DOCDB.
Domain: Up to 500 characters
Most names are only up to 30 characters in length.
Default value: n/a
Source database: DOCDB
Source field name

```
<applicants>
      <applicant sequence="1" data-format="docdb" status="A">
            <applicant-name>
                  <name>MACDONALD ALEX BRUCE</name>
            </applicant-name>
            <residence>
                  <country>US</country>
            </residence>
      </applicant>

<inventors>
      <inventor sequence="1" data-format="docdb">
            <inventor-name>
                  <name>MACDONALD ALEX BRUCE</name>
            </inventor-name>
            <residence>
                  <country>US</country>
            </residence>
</inventor>
```

Source sub-field identifier
`data-format="docdb"`
Comments
It is not 100% certain that the DOCDB standardised names are always linked with the correct person name, in particular if the person information came from a source other than DOCDB. This is especially true for names in USPTO patents. The reason is that the matching algorithm which merges the different sources relies that the names are being listed in the same sequence in all data sources (DOCDB and others), which is sometimes not the case.

In case DOCDB does not provide a DOCDB standardised name, this attribute will contain the same data as the attribute PERSON_NAME. Then and only then the attribute DOC_STD_NAME_ID will have a value > 100 000 000.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 15-04-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-10-2013 - Added comments about use in PATSTAT Online
M. Kracker - 15-10-2014 – Comment updated; comment added to domain
M. Kracker – 01-04-2015 – Comment amended
M. Kracker – 01-12-2015 – Length of domain extended
M. Kracker – 01-04-2014 – Comment amended

M. Kracker – 01-10-2017 – Comment amended to refer to STAN

6.45 DOC_STD_NAME_ID

Name: ID for the DOCDB standardised name
Also Known As: n/a
Description: DOC_STD_NAMEs which have been standardised according to the DOCDB standardisation procedure have a unique DOC_STD_NAME_ID for each unique DOC_STD_NAME. Multiple rows may have the same DOC_STD_NAME_ID, if multiple person names in the person table have been harmonised into a single DOCDB standard name.
DOC_STD_NAMEs which have *not* been standardised this way have a unique DOC_STD_NAME_ID for each (unstandardised) PERSON_NAME.
Domain: Number 1 … 999 999 999
Default value: n/a
Source database: PATSTAT
Source field name: Computed:
Not all DOC_STD_NAMEs have undergone the standardisation process.

- For DOC_STD_NAMEs which have been standardised the unique ID for each DOC_STD_NAME is in the range 1 ... 100 000 000
- For DOC_STD_NAMEs which have *not* been standardised but which just have been replenished by the PERSON_NAME the number is computed as “PERSON_ID + 100 000 000”.

Source sub-field identifier: n/a
Comments

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 15-04-2005 - First version
M. Kracker - 15-10-2014 – Comment amended
M. Kracker – 01-04-2015 – Stability assertion removed
M. Kracker - 01-04-2017 – Clarifications in Description, Source Database and Comment
M. Kracker - 01-10-2019 – Correction in the Source field name specification

6.46 DOCUS_CLASS_SYMBOL

Name: Symbols defined within the US Patent Classification (USPC) of the USPTO
Also Known As: n/a
Description: The US Patent Classification (USPC) is now superseded by CPC (Cooperative Patent Classification). USPC is now only used in special cases like for plant patents or design patents.
Domain: Up to 50 ASCII characters (as of 2019 Spring Edition, all symbols have a length between 8 and 10 characters)
Default value: n/a
Source database: DOCDB
Source field name

```
<patent-classifications>
      <patent-classification sequence="1">
            <classification-scheme office="US" scheme="DOCUS"/>
            <classification-symbol>081058000X</classification-symbol>
```

Source sub-field identifier
n/a
Comments
These classifications are being stored in DOCDB as supplied by the National Office without inspection of the contents. The EPO does not hold any responsibility for content, format or validity.

In many cases the content of DOCUS_CLASS_SYMBOL are structured like this:

			Examples	
Position	**Level**	**Comment**	**DOCUS_CLASS_SYMBOL**	**USPC**
1-3	class	leading zeroes; separator suppressed; may be alphanumerical	**101**401000P **D18**015000S **PLT**104000P	**101**/401 **D18**/15 **PLT**/104
4--6	subclass	leading zeroes; separator suppressed; may be alphanumerical	101**401**000P D18**015**000S 216**008**000X	101/**401** D18/**15** 216/**8**
7-9	extension	trailing zeroes when subclass is numerical leading zeroes when subclass is alphanumerical leading zeroes when extension is alphanumerical	123090**240**O 002DIG**003**S 134058**0DL**P	12390.**24** 2/DIG. **3** 134/58 **DL**
10	letter	‘S’ or ‘P’ when published in 1st document; O’ or ‘X’ when published in 2nd document; added by EPO at database loading time		

Modification history
Author of update - Date of update - Explanation of update
D. Lingua - 04-08-2011 - First version
M. Kracker - 01-10-2019 – Update of description and comment

6.47 EARLIEST_FILING_DATE

Name: Date of the earliest filing
Also Known As: n/a
Description: The earliest date of the filing dates of the application itself, its international application, its Paris Convention priority applications, the applications with which it is related via technical relations and its application continuations.
Only directly related applications are considered, e.g. not priorities of priorities.
Domain: Date (up to 9999-12-31)
Default value: 9999-12-31
Source database: PATSTAT
Source field name:
It is the APPLN_FILING_DATE of the earliest filing (see attribute EARLIEST_FILING_ID)

Source sub-field identifier: n/a

Comments:
For more details see EARLIEST_FILING_ID

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 15-10-2014 – Description clarified
M. Kracker - 01-10-2015 – Name of attribute has changed (was PRIOR_EARLIEST_DATE)

6.48 EARLIEST_FILING_ID

Name: Application ID of the earliest filing
Also Known As: First filing
Description: The ID of the earliest application, considering the application itself, its international application, its Paris Convention priority applications, the applications with which it is related via technical relations and its application continuations.
Only directly related applications are considered, e.g. not priorities of priorities.
Domain: Number 0 … 999 999 999;
Surrogate key: Technical unique identifier without any business meaning
Default value: n/a
Source database: PATSTAT
Source field name: Derived from the tables

- TLS201_APPLN self-priority
- TLS201_APPLN PCT application (= international application)
- TLS204_APPLN_PRIOR Paris Convention priority
- TLS205_TECH_REL technical relations
- TLS216_APPLN_CONTN application continuations

Source sub-field identifier: n/a
Comments:
If multiple applications have been filed on the earliest filing date, then conceptually any of these applications can be regarded as the earliest application. Nevertheless, the logic to determine the application which has been filed first is like this:

1. If there is a PCT application which was filed on the earliest application date, then the APPLN_ID of this PCT application is taken as the EARLIEST_FILING_ID.
2. Else: If there are 1 or more Paris convention priorities which were filed on the earliest application date, then the Paris convention priority with the smallest APPLN_ID is taken as the EARLIEST_FILING_ID.
3. Else: the application which was filed on the earliest application date with the smallest APPLN_ID will be taken.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2015 - First version
M. Kracker - 01-04-2016 – Detailed rules for selection of earliest filed application has been added to comment.

6.49 EARLIEST_FILING_YEAR

Name: Year of the earliest filing date
Also Known As: n/a
Description: Year of the earliest filing date
Domain: 4 digits in the form yyyy (e. g. 2015)
Default value: n/a
Source database: PATSTAT
Source field name: Derived from attribute EARLIEST_FILING_DATE of table TLS201_APPLN.
It is the year component of the attribute EARLIEST_FILING_DATE.

Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-10-2015 – Name of attribute has changed (was: PRIOR_EARLIEST_YEAR)

6.50 EARLIEST_PAT_PUBLN_ID

Name: ID of the earliest publication of an application
Also Known As: n/a
Description: The ID of a publication published on the earliest publication date of an application. Earlier applications, e.g. of the same patent family, are not considered.
Domain: Number 0 … 999 999 999
Default value: 0
Source database: PATSTAT
Source field name: The EARLIEST_PAT_PUBLN_ID is the PAT_PUBLN_ID from table TLS211_PAT_PUBLN whose publication has been published on the date EARLIEST_PUBLN_DATE in table TLS201_APPLN of its application.

Source sub-field identifier: n/a
Comments: If more than one publication is published on the same (earliest) publication date, then any one is selected.
All publications of table TLS211_PAT_PUBLN are considered when computing this attribute. This includes publications which are only announcements in the Gazette of a patent office or which are only "laid open to the public", and which are in some contexts not regarded as formal publications. An example of such a publication is GB 0329008 D0 (PAT_PUBLN_ID = 424991854), which was published several months before the A-publication of its application.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2017 – Description and Comment amended
M. Kracker - 01-04-2020 – Correction and clarification in the source field name

6.51 EARLIEST_PUBLN_DATE

Name: Date of earliest publication of an application
Also Known As: n/a
Description: Date of earliest publication of an application. Earlier applications, e.g. of the same patent family, are not considered.
Domain: Date (up to 9999-12-31)
Default value: 9999-12-31
Source database: PATSTAT
Source field name: Derived from table TLS211_PAT_PUBLN.

It is the earliest PUBLN_DATE of the publications identified by TLS211_PAT_PUBLN.APPLN_ID = TLS201_APPLN.APPLN_ID .

Source sub-field identifier: n/a
Comments:
All publications of table TLS211_PAT_PUBLN are considered when computing this attribute. This includes publications which are only announcements in the Gazette of a patent office or which are only "laid open to the public", and which are in some contexts not regarded as formal publications. An example of such a publication is GB 0329008 D0 (PAT_PUBLN_ID = 424991854), which was published several months before the A-publication of its application.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-10-2015 - Name of attribute has changed (was: PUBLN_EARLIEST_DATE)
M. Kracker - 01-04-2017 - Description amended, Comment added
M. Kracker - 01-04-2020 - Correction and clarification in the source field name

6.52 EARLIEST_PUBLN_YEAR

Name: Year of the earliest publication date of an application
Also Known As: n/a
Description: Year of the earliest publication date of an application. Earlier applications, e.g. of the same patent family, are not considered.
Domain: 4 digits in the form yyyy (e. g. 2015)
Default value: n/a
Source database: PATSTAT
Source field name: Derived from attribute EARLIEST_PUBLN_DATE of table TLS201_APPLN ;
It is the year component of the attribute EARLIEST_PUBLN_DATE.
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2015 – Computation explained
M. Kracker - 01-10-2015 – Name of attribute has changed (was: PUBLN_EARLIEST_YEAR)
M. Kracker - 01-04-2017 – Clarification in description

6.53 EPO_MEMBER

Name: Member of the European Patent Organisation
Also Known As: n/a
Description: Indicates whether this country/territory is a member state of the EPO
Domain: 1 ASCII character: Y or space

Y	If a country/territory is member of the EPO. Only full members are considered, no contracting states or extension states.
space	otherwise

Default value: n/a
Source database: https://www.epo.org/about-us/organisation/member-states.html
Source field name: n/a
Source sub-field identifier: n/a
Comments: This field indicates the EPO members at the time of the production of the PATSTAT edition. Depending on the time range you need to analyse, you may want to exclude “newer” EPO members (e.g. AL, RS).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version

6.54 EU_MEMBER

Name: Member of the European Community
Also Known As: n/a
Description: Indicates whether this country/territory is a member state of the European Union
Domain: 1 ASCII character: Y or space

Y	If a country/territory is a member of the EU
space	otherwise

Default value: n/a
Source database: http://europa.eu/abc/european_countries/eu_members/index_de.htm
Source field name: n/a
Source sub-field identifier: n/a
Comments: This field indicates the EU members at the time of the production of the PATSTAT edition. Depending on the time range you need to analyse, you may want to exclude “newer” EU members (e.g. HR, BU, RO) or include former EU members (GB).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version

6.55 EVENT_AUTH

Name: Event authority
Also Known As: n/a
Description: The national office which has provided the legal event.
Domain: 2 ASCII characters (A-Z), according to WIPO ST.3
Default value: n/a

Source database: INPADOC (EPO worldwide legal status database)
Source field name: /legal-status-document/legal-event/@providing_office

```
<legal-event event-type="REG" providing-office="GB" date-added="20111103" date-
previous-exchange="20111103" sequence-number="11">
     <event-date>20111102</event-date>
     <event-code>FG4D</event-code>
     <event-details>
          <event-description event-description-type="original">EUROPEAN PATENT
GRANTED</event-description>
          <event-description lang="en">EUROPEAN PATENT GRANTED</event-description>
          <event-description lang="de">EUROPAEISCHES PATENT ERTEILT</event-
description>
          <event-description lang="fr">BREVET EUROPEEN DELIVRE</event-description>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.56 EVENT_CATEGORY_CODE

Name: Code of legal event category
Also Known As: n/a
Description: Code of legal event category
Domain: 1 ASCII character: A-Z or space
Default value: space

Source database: Based on Excel table “Legal status codes”, accessible on https://www.epo.org/searching-for-patents/data/coverage/weekly.html
Source field name: Column “Event-class”

Comments
INPADOC legal event codes are assigned by the EPO to categories of the INPADOC classification scheme. These categories are largely aligned with the categories of WIPO standard ST.27 "Recommendation for the exchange of patent legal status data” (http://www.wipo.int/export/sites/www/standards/en/pdf/03-27-01.pdf). However, they may differ if the primary objective of the INPADOC classification scheme, which is to help patent information users to understand and retrieve INPADOC legal event data, requires it.

More information on the category level of the INPADOC classification scheme can be found in the manual “INPADOC classification scheme” downloadable from https://www.epo.org/searching-for-patents/data/bulk-data-sets/manuals.html.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2018 - First version

6.57 EVENT_CATEGORY_TITLE

Name: Title of legal event category
Also Known As: n/a
Description: Label of legal event category
Domain: Up to 50 ASCII characters
Default value: n/a

Source database: Based on Excel table “Legal status codes”, accessible on https://www.epo.org/searching-for-patents/data/coverage/weekly.html
Source field name: Column “Event-class Description”

Comments
For details and explanation see attribute EVENT_CATEGORY_CODE.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2018 - First version

6.58 EVENT_CODE

Name: Legal event code
Also Known As: n/a
Description: The code which - in conjunction with the country code of the application - uniquely identifies a legal event.
Domain: Up to 4 ASCII characters
Default value: n/a

Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-code

```
<legal-event event-type="REG" providing-office="GB" date-added="20111103" date-
previous-exchange="20111103" sequence-number="11">
     <event-date>20111102</event-date>
     <event-code>FG4D</event-code>
     <event-details>
          <event-description event-description-type="original">EUROPEAN PATENT
GRANTED</event-description>
          <event-description lang="en">EUROPEAN PATENT GRANTED</event-description>
          <event-description lang="de">EUROPAEISCHES PATENT ERTEILT</event-
description>
          <event-description lang="fr">BREVET EUROPEEN DELIVRE</event-description>
     </event-details>
</legal-event>
```

Comments
INPADOC (EPO worldwide legal event database) uses a few thousand codes to classify legal events in the lives of industrial property rights. The content and coverage of the INPADOC legal event file can be found in https://www.epo.org/searching-for-patents/data/coverage/weekly.html.

Always check which national patent law is used with the legal event code.

A full list of event codes is given in the documents "Legal status codes" on https://www.epo.org/searching-for-patents/data/coverage/weekly.html .
For example EVENT_AUTH = "AT" and EVENT_CODE = "ELJ" means "Ceased due to non-payment of the annual renewal fee in Austria".

3 event codes indicate a specific legal event of an EP patent, where the event actually took place in the *national phase* of these EP applications. The national office where this event takes place is indicated in attributes FEE_COUNTRY resp. LAPSE_COUNTRY resp. REINSTATE_COUNTRY. The 3 event codes are:

- PGFP Post grant: Annual fees paid to the national office
- PG25 Lapsed in a contracting state announced via post grant information from national office to EPO
- PGRI Post grant: Patent reinstated in contracting state

Each of these codes has additional attributes in TLS231_INPADOC_LEGAL_EVENT which give more information.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.59 EVENT_DESCR

Name: Description of the legal event code in English
Also Known As: n/a

Description: Short english text explaining the legal event code
Domain: Up to 250 characters
Default value: n/a
Source database: Based on Excel table “Legal status codes”, accessible on the EPO homepage at https://www.epo.org/searching-for-patents/data/coverage/weekly.html
Source field name: column “Description ENG”

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.60 EVENT_DESCR_ORIG

Name: Description of the legal event code in the original language
Also Known As: n/a

Description: Short english in original language explaining the legal event code.If the original language is not available, the description will be in English.
Domain: Up to 250 characters
Default value: n/a
Source database: Based on Excel table “Legal status codes”, accessible on the EPO homepage at https://www.epo.org/searching-for-patents/data/coverage/weekly.html
Source field name: column “Description ORI”

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.61 EVENT_EFFECTIVE_DATE

Name: Effective date
Also Known As: n/a
Description: The date this event has come into force.
Domain: Date
Default value: 9999-12-31
Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-date-effective

```
<legal-event event-type="REG" providing-office="DE" date-added="20120102" date-
previous-exchange="20120105" sequence-number="14">
     <event-date>20111229</event-date>
     <event-date-effective>20111229</event-date-effective>
     <event-code>R096</event-code>
     <event-details>
          <event-description event-description-type="original">VEROEFFENTLICHUNG
EINES HINWEISES AUF DIE EP-PATENTERTEILUNG DURCH DAS DPMA</event-description>
          <event-description lang="en">DPMA PUBLICATION OF MENTIONED EP PATENT
GRANT</event-description>
          <event-reference>
               <event-ref-document>
                    <country>DE</country>
                    <doc-number>602010000345</doc-number>
               </event-ref-document>
          </event-reference>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.62 EVENT_FILING_DATE

Name: Event filing date
Also Known As: n/a
Description: The date the event has been filed. Note that this attribute is rarely populated. It is often related to patent term extensions, re-examination decisions or limitations.
Domain: Date
Default value: 9999-12-31
Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-details/date-filing

```
<legal-event providing-office="US" date-added="20160210" date-previous-
exchange="20160416" sequence-number="6">
      <event-date>20160209</event-date>
      <event-date-effective>20160204</event-date-effective>
      <event-code>LIMR</event-code>
      <event-details>
            <event-description lang="en">REEXAMINATION DECISION: CLAIMS CHANGED
AND/OR CANCELLED</event-description>
            <date-filing>20140829</date-filing>
            <text>CLAIMS 1-8, 13 AND 18-26 ARE CANCELLED. CLAIMS 9-12 AND 14-17
WERE NOT REEXAMINED.</text>
      </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2017 – New attribute

6.63 EVENT_ID

Name: Identifier for a legal event
Also Known As: n/a
Description: Technical unique identifier for an INPADOC legal event
Domain: number
Default value: n/a
Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event@event-id

```
<legal-event providing-office="EP" event-id="18846513" date-added="20030128"
date-previous-exchange="20030101" sequence-number="2">
      <event-date>19900627</event-date>
      <event-date-effective>19900425</event-date-effective>
      <event-code>17P</event-code>
      <event-details>
            <event-description event-description-type="original">PRUEFUNGSANTRAG
GESTELLT</event-description>
            <event-description Lang="en">REQUEST FOR EXAMINATION FILED</event-
description>
      </event-details>
</legal-event>
```

Source sub-field identifier: n/a

Comments
This is a stable attribute, which means that its value will not change between PATSTAT editions.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2018 – New attribute

6.64 EVENT_IMPACT (deprecated)

Name: Impact of the legal event on the patent
Also Known As: n/a

Description: Whether the legal event is perceived to have a positive, negative or neutral effect on the life of the patent.This attribute is an educated guess done by the EPO and has no legal basis.
Domain: 1 ASCII character: +, - or space

+	positive impact
-	negative impact
space	no impact

Default value: space
Source database: Based on Excel table “Legal status codes”, accessible on the EPO homepage https://www.epo.org/searching-for-patents/data/coverage/weekly.html .
Source field name: column “Influence” in Sheet “Codes”

Comments
This attribute is deprecated, because the source data will cease to be available in the near future. The attribute will be removed in one of the next PATSTAT editions.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version
M. Kracker - 01-10-2018 – Attribute is deprecated

6.65 EVENT_PUBLN_DATE

Name: Publication date of the legal event
Also Known As: n/a
Description: Date on which the legal event has been made available to the public, e. g. in a gazette or online
Domain: Date
Default value: n/a
Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-date

```
<legal-event event-type="REG" providing-office="GB" date-added="20111103" date-
previous-exchange="20111103" sequence-number="11">
     <event-date>20111102</event-date>
     <event-code>FG4D</event-code>
     <event-details>
          <event-description event-description-type="original">EUROPEAN PATENT
GRANTED</event-description>
          <event-description lang="en">EUROPEAN PATENT GRANTED</event-description>
          <event-description lang="de">EUROPAEISCHES PATENT ERTEILT</event-
description>
          <event-description lang="fr">BREVET EUROPEEN DELIVRE</event-description>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.66 EVENT_SEQ_NR

Name: Sequence number of the legal event
Also Known As: n/a
Description: Number indicating the sequence of the legal event in the series of legal events for one patent application
Domain: Number 1 … about 2.000; max value in 2017 Autumn Edition was 1.055
Default value: n/a
Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/@sequence-number

```
<legal-event event-type="REG" providing-office="GB" date-added="20111103" date-
previous-exchange="20111103" sequence-number="11">
     <event-date>20111102</event-date>
     <event-code>FG4D</event-code>
     <event-details>
          <event-description event-description-type="original">EUROPEAN PATENT
GRANTED</event-description>
          <event-description lang="en">EUROPEAN PATENT GRANTED</event-description>
          <event-description lang="de">EUROPAEISCHES PATENT ERTEILT</event-
description>
          <event-description lang="fr">BREVET EUROPEEN DELIVRE</event-description>
     </event-details>
</legal-event>
```

Comments
For a given patent application, each legal event is assigned a sequence number so as to make each row identifiable by the combination APPLN_ID and EVENT_SEQ_NR.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.67 EVENT_TEXT

Name: Additional information
Also Known As: n/a
Description: Additional information in free form text relating to an event which is not covered by another attribute.
Domain: Up to 1 000 characters
Default value: empty
Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-details/text

```
<legal-event event-type="REG" providing-office="GB" date-added="20121004" date-
previous-exchange="20121004" sequence-number="38">
     <event-date>20121003</event-date>
     <event-code>732E</event-code>
     <event-details>
          <event-description event-description-type="original">AMENDMENTS TO THE
REGISTER IN RESPECT OF CHANGES OF NAME OR CHANGES AFFECTING RIGHTS (SECT.
32/1977)</event-description>
          <event-description lang="en">AMENDMENTS TO THE REGISTER IN RESPECT OF
CHANGES OF NAME OR CHANGES AFFECTING RIGHTS (SECT. 32/1977)</event-description>
          <text>REGISTERED BETWEEN 20120906 AND 20120912</text>
     </event-details>
</legal-event>
```

Comments
Currently (Sept. 2016) the maximum length of this attribute is 700 characters.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.68 EVENT_TYPE

Name: Indicates whether an event refers to an international or regional application
Also Known As: n/a
Description: A value "REG" indicates that an event refers to the national or regional phase of an international or regional application. The patent authority which triggered the event is indicated in the attribute EVENT_AUTH (except for EVENT_CODEs PGFP, PG25 and PGRI; see 6.58 “EVENT_CODE” for more details)

Domain: up to 3 ASCII characters: "REG" or empty
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/@event-type

```
<legal-event event-type="REG" providing-office="GB" date-added="20111103" date-
previous-exchange="20111103" sequence-number="11">
     <event-date>20111102</event-date>
     <event-code>FG4D</event-code>
     <event-details>
          <event-description event-description-type="original">EUROPEAN PATENT
GRANTED</event-description>
          <event-description lang="en">EUROPEAN PATENT GRANTED</event-description>
          <event-description lang="de">EUROPAEISCHES PATENT ERTEILT</event-
description>
          <event-description lang="fr">BREVET EUROPEEN DELIVRE</event-description>
     </event-details>
</legal-event>
```

Comments
EVENT_TYPE = REG indicates that a legal event took place in a national phase of a regional or international application. It is provided by the national office. The REG code is not limited to EP applications, but is applied to these applications and their national phases:

- National phase of an EP application
- National phase of an PCT application
- RU (Russian) phase of an earlier SU (Soviet Union) application
- HK (Hong Kong) phase of an GB (United Kingdom) application
- HK (Hong Kong) phase of an CN (Chinese) application

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.69 EXTENSION_STATES

Name: Extension state(s)
Also Known As: n/a
Description: List of country codes of extension states.
Regional groupings of national patent offices exist such as the EPO. These regional offices sometimes allow applicants to extend protection to non-member states.
Domain: up to 30 ASCII characters, consisting of an alphabetically ordered list of 2-character country codes (according to WIPO ST.3), with each country separated by a comma ",":
Examples: "ME" or "BA,ME"
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/extension-states/country

```
<legal-event providing-office="EP" date-added="20101007" date-previous-
exchange="20101007" sequence-number="2">
     <event-date>20101006</event-date>
     <event-code>AX</event-code>
     <event-details>
          <event-description event-description-type="original">REQUEST FOR
EXTENSION OF THE EUROPEAN PATENT TO</event-description>
          <event-description lang="en">REQUEST FOR EXTENSION OF THE EUROPEAN
PATENT TO</event-description>
          <extension-states>
               <country>BA</country>
               <country>ME</country>
          </extension-states>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.70 FEE_COUNTRY

Name: Country or territory which received fee payment
Also Known As: n/a
Description: The office which received the payment of the annual renewal fee for a patent.
Domain: 2 ASCII characters (A-Z), according to WIPO ST.3
Default value: empty if EVENT_CODE is not “PGFP” or “VSFP”

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/fee-payment/@country

```
<legal-event providing-office="EP" date-added="20150804" date-previous-
exchange="20150808" sequence-number="49">
     <event-date>20150731</event-date>
     <event-code>PGFP</event-code>
     <event-details>
          <event-description event-description-type="original">POSTGRANT: ANNUAL
FEES PAID TO NATIONAL OFFICE</event-description>
          <event-description lang="en">POSTGRANT: ANNUAL FEES PAID TO NATIONAL
OFFICE</event-description>
          <fee-payment country="DE">
               <fee-payment-date>20150331</fee-payment-date>
               <fee-payment-year>6</fee-payment-year>
          </fee-payment>
     </event-details>
</legal-event>
```

Comments
Note that for EP patents the annual renewal fees are not paid to the EPO but to the respective National Office of the EPO member state for which the patent should be renewed.

This attribute is populated if and only if EVENT_CODE = "PGFP" (Post Grant Fee Paid for an EP patent) or “VSFP” (Annual fee paid to validation state). For PGFP see also Business Rules about table TLS231_INPADOC_LEGAL_EVENT in section5.24.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version
M. Kracker - 01-04-2021 – Will also be populated for EVENT_CODE = VSFP

6.71 FEE_PAYMENT_DATE

Name: Fee payment date
Also Known As: n/a
Description: Date of payment of the annual renewal fee for an EP patent
Domain: Date
Default value: 9999-12-31

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/fee-payment/fee-payment-date

```
<legal-event providing-office="EP" date-added="20150804" date-previous-
exchange="20150808" sequence-number="49">
     <event-date>20150731</event-date>
     <event-code>PGFP</event-code>
     <event-details>
          <event-description event-description-type="original">POSTGRANT: ANNUAL
FEES PAID TO NATIONAL OFFICE</event-description>
          <event-description lang="en">POSTGRANT: ANNUAL FEES PAID TO NATIONAL
OFFICE</event-description>
          <fee-payment country="DE">
               <fee-payment-date>20150331</fee-payment-date>
               <fee-payment-year>6</fee-payment-year>
          </fee-payment>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.72 FEE_RENEWAL_YEAR

Name: Year of annual renewal fee payment
Also Known As: n/a
Description: Year of annual renewal fee payment for an EP patent, e. g. 7 for the seventh year.
Domain: number 1 .. 25 (typically up to 20; exceptions are SPCs, etc.)
Default value: 0

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/fee-payment/fee-payment-year

```
<legal-event providing-office="EP" date-added="20150804" date-previous-
exchange="20150808" sequence-number="49">
     <event-date>20150731</event-date>
     <event-code>PGFP</event-code>
     <event-details>
          <event-description event-description-type="original">POSTGRANT: ANNUAL
FEES PAID TO NATIONAL OFFICE</event-description>
          <event-description lang="en">POSTGRANT: ANNUAL FEES PAID TO NATIONAL
OFFICE</event-description>
          <fee-payment country="DE">
               <fee-payment-date>20150331</fee-payment-date>
               <fee-payment-year>6</fee-payment-year>
          </fee-payment>
     </event-details>
</legal-event>
```

Comments
Only the most recent payment is recorded.

This attribute must not be confused with attribute FEE_PAYMENT_DATE. Example, assuming the last payment, made on the 01. 04. 2020, was for the 7th renewal fee, then FEE_PAYMENT_YEAR will be "7" (and not "2020").

Not all Offices request a payment for an annual renewal fee for each operating year (OY). Exceptions as of 2018 are:

- **Netherlands** have no fee for OY3.
- **Switzerland** has no fee for OY3.
- **Great Britain** has no fee for OY3 and 4.
- **Austria** has no fee for OY 3, 4 and 5.
- **San Marino** has no fee in year 3, but since they require a translation to be filed, there are PGFP's for OY3 if needed in the year of due payment
- **Italy** has no fee in years 3 or 4, but since they require a translation to be filed, there are PGFP's for OY3 & 4 if needed in the year of due payment

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

M. Kracker - 01-10-2018 – Comment added to list exception to annual payments for EP member states

6.73 FEE_TEXT

Name: Additional information about a payment
Also Known As: n/a
Description: Additional information in free form text about the annual renewal fee payment for an EP patent.
Domain: up to 1 000 characters
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/fee-payment/text

```
<legal-event providing-office="US" date-added="20131127" date-previous-
exchange="20131130" sequence-number="1">
     <event-date>19900706</event-date>
     <event-code>FPAY</event-code>
     <event-details>
          <event-description event-description-type="original">FEE PAYMENT</event-
description>
          <event-description lang="en">FEE PAYMENT</event-description>
          <fee-payment>
               <fee-payment-date>19900706</fee-payment-date>
               <fee-payment-year>4</fee-payment-year>
               <text>PAYMENT OF MAINTENANCE FEE, 4TH YEAR, PL 97-247</text>
          </fee-payment>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.74 FIRST_NAME

Name: First name of a physical person
Also Known As: n/a
Description: Contains the first name of a physical person
Domain: Up to 500 characters
Default value: empty string
Source database: USPTO data of published applications and published grants
Source field name:
`<addressbook><first-name>`
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2016 – Domain extended to 500 characters

6.75 GRANTED

Name: "Granted" indicator
Also Known As: n/a
Description: Y if this application has been granted; N otherwise
Domain: 1 ASCII character: Y or N
N – there is *no* indication in the data that the application has been granted
Y – there is an indication in the data that the application has been granted
Default value: n/a
Source database: PATSTAT
Source field name: Derived from attribute PUBLN_FIRST_GRANT of table TLS211_PAT_PUBLN and from legal events in table TLS231_INPADOC_LEGAL_EVENT:

If the application has a publication with PUBLN_FIRST_GRANT = “Y” or / and the application has a legal event which is in a legal category with EVENT_CATEGORY_CODE = “F” (which means “IP RIGHT GRANT”), then GRANTED will have the value “Y”. Otherwise GRANTED will have the value “N”.

Note: For international applications (APPLN_AUTH = ‘WO’) “granted” means that the application was granted in one or more of its designated states. The detailed information which designated states have in this way “granted” an international application can be retrieved from table TLS231_INPADOC_LEGAL_EVENT.
Example: APPLN_AUTH = WO and APPLN_NR = 9919007 has been granted in Australia.

Source sub-field identifier: n/a
Comments:
Some offices, e.g. Argentina, Brazil or Mexico, do not (always) publish granted patents but just issue a legal event. So the legal events in table TLS231_INPADOC_LEGAL_EVENT are utilized to reveal additional indications for grants

Although the EPO has taken great care in analysing the grant information, this process is the result of interpretations and assumptions for which no responsibility whatsoever can be accepted.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2014 – Comment added
M. Kracker - 01-10-2014 – Comment amended
M. Kracker - 01-10-2018 – Domain changed from 0/1 to N/Y; Computation – and therefore the content – of the attribute has changed; Comment has been adapted.

6.76 HAN_HARMONIZED

Name: Harmonisation indicator for OECD HAN
Also Known As: n/a
Description: Indicates the degree of harmonisation and standardisation which could be achieved
Domain: Number 0 … 2

0	the HAN_NAME has been replenished with the original name, because the name could not be harmonised.
1	the HAN_NAME has been harmonised but could not be matched with the ORBIS© database.
2	the HAN_NAME has been harmonised and could be matched with the ORBIS© database.

Default value: n/a
Source database: OECD HAN database
Source field name: n/a
Source sub-field identifier: n/a
Comments:
The processing of the Harmonised Applicant Name starts as soon as PATSTAT data is released. Typically the result will be available 3-4 months afterwards. Consequently, additions and changes introduced by the current PATSTAT edition are not harmonised in the current edition but will be harmonised in the next edition of PATSTAT.

See also comment of attribute HAN_NAME.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 15-10-2014 –Comment updated
M. Kracker - 01-04-2016 –Comment updated

6.77 HAN_ID

Name: ID of a Harmonised Applicant Name (HAN) from OECD
Also Known As: OECD HAN ID, HAN ID
Description:
HAN_NAMEs which have been harmonised according to the OECD harmonisation procedure have a unique HAN_ID for each unique HAN_NAME. Multiple rows may have the same HAN_ID, if multiple person names in the person table have been harmonised into a single HAN name.
HAN_NAMEs which have *not* been harmonised this way have a unique HAN_ID for each (un-harmonised) PERSON_NAME.
Domain: Number 1 … 999 999 999
Default value: n/a
Source database: Computed from OECD HAN database
Not all HAN_NAMEs have undergone the harmonisation process (cf. attribute HAN_HAMONIZED).

- For HAN_NAMEs which have *not* been created during the harmonisation process the unique HAN_ID for each HAN_NAME is in the range 1 ... 100 000 000
- For HAN_NAMEs which have *not* been created during the harmonisation process but which just have been replenished by the PERSON_NAME the number is computed as “PERSON_ID + 100 000 000”.

Source field name: n/a
Source sub-field identifier: n/a
Comments:

The processing of the Harmonised Applicant Name starts as soon as PATSTAT data is released. Typically the result will be available 3-4 months afterwards. Consequently, additions and changes introduced by the current PATSTAT edition are not harmonised in the current edition but will be harmonised in the next edition of PATSTAT.

See also comment of attribute HAN_NAME.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 15-10-2015 – Comment updated
M. Kracker - 01-04-2015 – Comment updated
M. Kracker - 01-04-2016 – Comment updated
M. Kracker - 01-04-2017 – Clarifications in Description and Source Database

6.78 HAN_NAME

Name: Harmonised Applicant Name (HAN) from OECD
Also Known As: OECD HAN name, HAN name
Description: This field contains for many applicants the names as harmonised by the OECD HAN (Harmonised Applicant Name) project of the OECD.

The scope of this harmonisation effort is described by the OECD as:
"The OECD HAN database, July 2014, provides groupings of patent applicant's names for the following set of countries or economies : {AR. AT, AU, BE, BR, CA, CH, CL, CN, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IL, IN, IS, IT, JP, KR, LU, MX, NL, NO, NZ, PL, PT, RU, SE, SI, SK, TR, TW, US, ZA}. The list of patents filed to the EPO, the USPTO and through the PCT is made available for each grouping of applicants. Further improvements are expected in future version, notably on the countries coverage."

The attribute is populated for all persons. Names of persons which have not been harmonised (e. g. persons who are inventors but not applicants) are just copied from the attribute PERSON_NAME.
Domain: Up to 500 characters
Default value: n/a
Source database: OECD HAN database
Source field name: n/a
Source sub-field identifier: n/a
Comments:
These names have been taken from the OECD HAN database (cf. http://www.oecd.org/sti/innovationinsciencetechnologyandindustry/oecdpatentdatabases.htm).

The processing of the Harmonised Applicant Name starts as soon as PATSTAT data is released. Typically the result will be available 3-4 months afterwards. Consequently, additions and changes introduced by the current PATSTAT edition are not harmonised in the current edition but will be harmonised in the next edition of PATSTAT.

Please note that the OECD HAN database is provided for research and analytical work. When publishing the results of your analysis, make sure it is quoted as: "OECD, HAN database, <Month, Year>".

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 15-10-2014 – Update of description and comment
M. Kracker - 01-04-2016 –Comment updated

6.79 INPADOC_FAMILY_ID

Name: Identifier of an INPADOC extended priority family
Also Known As: INPADOC family ID; Extended family ID
Description: Means that the applications share a priority directly or indirectly via a third application. A 'priority' in this case means a link shown between applications as in tables TLS201_APPLN (regional/national phase of a PCT application), TLS204_ APPLN_PRIOR (PARIS convention priorities), TLS205_TECH_REL (patents which have been technically linked by patent examiners on the basis of similar content) and table TLS216_ APPLN_CONTN (continuations, divisions etc.).

Domain: Number 0 … 999 999 999

Default value: n/a
Source database: This attribute is calculated during the preparation of PATSTAT data. For the dummy application (i.e. APPLN_ID = 0) and for artificial application replenished because of citations (i.e. APPLN_ID > 930 000 000) the value of the INPADOC_FAMILY_ID will be the same as the value of the APPLN_ID.
Source field name: n/a
Source sub-field identifier: n/a

Comments
Every application belongs to exactly one INPADOC family. In the trivial case, an application belongs to an INPADOC family which consists of no other family members except this application itself. This is e.g. the case for all artificial applications with an APPLN_ID > 930 000 000; see section 4.4).

Much patent research is affected by the “family” concepts. There are various definitions of how to link different patents into “families”. This INPADOC extended priority family was developed by the INPADOC organisation before it was integrated into the EPO.

For the computation of the INPADOC families these tables are considered:

- TLS201_APPLN
  A PCT application in its regional/national phase contains in its attribute INTERNAT_APPLN_ID the APPLN_ID of its original PCT application
- TLS204_appln_prior (Paris convention priorities)
- TLS205_TECH_REL (patents which have been technically linked by patent examiners on the basis of similar content)
- TLS216_appln_contn (continuations, divisions, etc.).

The EPO reserves the right to apply this rule as needed for its internal purposes.

The simplified definition of the INPADOC family is that the members relate in some way (directly or indirectly) to the "first" application. Generally an INPADOC family covers one or more DOCDB families and covers a set of related inventions.

From a statistical point of view: a large DOCDB family might indicate that the applicant seeks a wide geographical protection for the invention.

The value of the INPADOC_FAMILY_ID is not stable but will change with every edition of PATSTAT. For technical reasons, the INPADOC_FAMILY_ID will be identical to the smallest APPLN_ID of all members of that INPADOC family.

Modification history
Author of update - Date of update - Explanation of update
J. Rollinson / D. Lingua - 19-09-2008 - First version
D. Lingua - 11-10-2011 - Comment on SQL queries eliminated
M. Kracker - 01-10-2013 - Change of domain
M. Kracker - 01-10-2015 – Value 0 added to domain; Change of comment and description
M. Kracker - 01-10-2016 – Change of comment
M. Kracker - 01-10-2019 – Change of comment

6.80 INT_PHASE

Name: Indicator whether the application *is* or *has been* in the international phase
Also Known As: n/a
Description: Indicates that an application *is* or *has been* in the international phase. This covers all international filings as well as all applications based on these filings.
Domain: 1 ASCII character

Y	Yes
N	No
space	not known (In case of uncertain interpretations; used very little or not at all)

Default value: N
Source database: PATSTAT
Source field name: Derived from table TLS201_APPLN

Y	if the application has APPLN_KIND = W (i.e. international filing) or INTERNAT_APPLN_ID > 0; (i.e. based on internat. application)
N	otherwise

Source sub-field identifier: n/a
Comments:
These indicators provide a somewhat simplistic approach to identify the route an application has taken. Although the EPO has taken great care in analysing the underlying data (especially publication and application kind codes), this process is the result of interpretations and assumptions for which no responsibility whatsoever can be accepted.

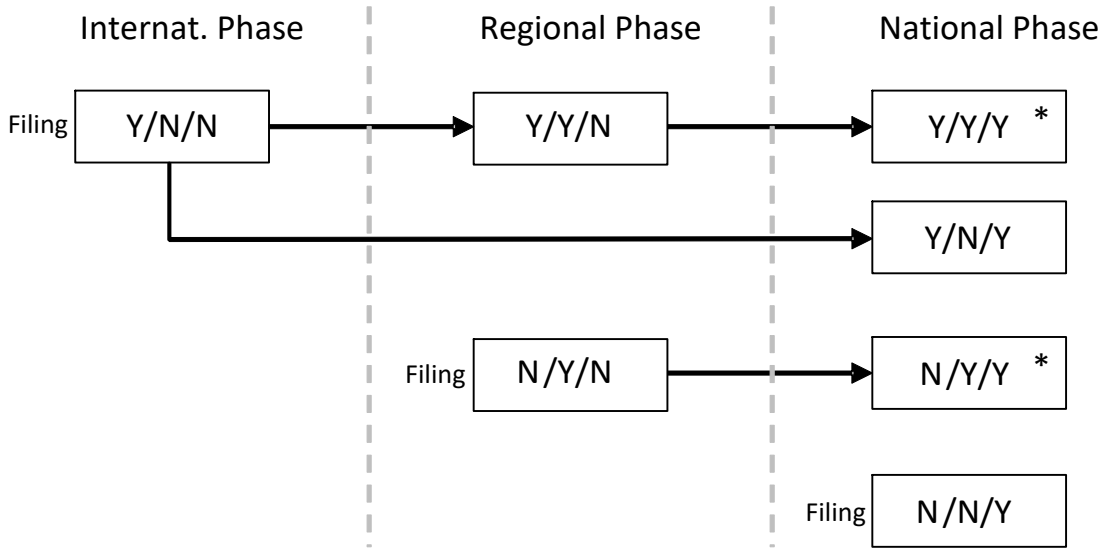
Please note that these indicators only help to understand applications which actually exist in PATSTAT. It does not help to answer questions like “How many EP applications are valid in country x”, because not every office publishes patents which are validated / granted in their country. Consequently, there is no publication or application in PATSTAT for every granted patent.
The same will apply for the Unitary Patents, if there is no publication for that.

Especially if you want to count the patents of an office which is a member of a regional office: As argued above, it is not sufficient to consider the indicator NAT_PHASE. Depending on the office, you must also analyse the legal events of the application in the regional phase (see also table TLS231_INPADOC_LEGAL_EVENT or the database ”PATSTAT EP Register” for the legal events of EP applications.

An application can take one of these routes:

Possible routes of an application

Boxes contain the values of attributes INT_PHASE / REG_PHASE / NAT_PHASE



* Often these regional patents in the national phase are not published and therefore they are not in PATSTAT.

	Values in attributes		
Application ...	INT_PHASE	REG_PHASE	NAT_PHASE
... in the international phase	Y	N	N
... PCT in the regional phase (e. g. Euro-PCT)	Y	Y	N
... PCT via regional office, now in national phase	Y	Y	Y
... PCT in the national phase (no regional phase)	Y	N	Y
... in the regional phase (no PCT)	N	Y	N
... via regional office, now in national phase (no PCT)	N	Y	Y
... national application (no PCT)	N	N	Y

Modification history

Author of update - Date of update - Explanation of update

M. Kracker - 01-04-2016 - First version

6.81 INTERNAT_APPLN_ID

Name: Application identification of the earlier PCT international application for an application.
Also Known As: n/a
Description: Technical unique identifier without any business meaning
Domain: Number 0 … 999 999 999
Default value: 0
Source database: DOCDB, PATSTAT
Source field name
International applications designating the Authority of the related national / regional application. The **latter** is published with an INID-code in the 80-series (WIPO ST.9). The relevant case is case # 3 from section 4.6 "Relation Types".

```
<priority-claims>
      <priority-claim sequence="1" data-format="docdb">
            <document-id>
                  <country>US</country>
                  <doc-number>0107931</doc-number>
                  <kind>W</kind>
                  <date>20010312</date>
            </document-id>
            <priority-linkage-type>W</priority-linkage-type>
      <priority-active-indicator>N</priority-active-indicator>
</priority-claim>
```

With

```
<document-id>
      <country>US</country>
      <doc-number>0107931</doc-number>
      <kind>W</kind>
```

in DOCDB the corresponding international application in PATSTAT is determined (via APPLN_AUTH, APPLN_NR and APPLN_KIND) and the value of INTERNAT_APPLN_ID for this national/regional application is set to the value of APPLN_ID of the international application. If there is no corresponding international application in PATSTAT it should be created, see section 4.4 "Application replenishment".

Source sub-field identifier
n/a

Comments
The default value 0 means this application has no earlier PCT application.
If the value of INTERNAT_APPLN_ID is > 0, then this application does have an earlier PCT application, whose APPLN_ID equals the value of INTERNAT_APPLN_ID.

Note that for some countries there will be no applications with INTERNAT_APPLN_ID >0, because for these countries the “national route” via the PCT has been closed (for a list of these countries see http://www.wipo.int/export/sites/www/pct/en/list_states.pdf, footnote 2). For example, France does not accept PCT applications to go directly to France. Instead, for the PCT application the appropriate regional office (for FR it is EP) must be designated first, and the granted regional patent may then be validated in France.

Modification history

Author of update - Date of update - Explanation of update
R. Heijna - 03-05-2005 - First version
R. Heijna - 20-07-2005 - Source field definition improved
R. Heijna - 07-07-2005 - Value zero for the physical model
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 27-07-2010 - Revision of text
D. Lingua - 17-04-2011 - Warning added
M. Kracker - 10-04-2015 –Added comment: for some countries the national phase via the PCT is closed
M. Kracker - 01-04-2019 –Removed rule introduced on 08-10-2012 to identify Euro-PCTs

6.82 INVT_SEQ_NR

Name: Sequence number of inventors
Also Known As: n/a
Description: Number indicating the place in the list of inventors in the application
Domain: Number 0 … about 250
Default value: 0
Source database:
1) EP Register for EP patent applications
Contains the sequence numbers.

2) OECD patents database for US data post 1976-01-01 up to and including November 15th, 2005 for Published Grants. This data does not contain sequence numbers, so they are allocated within PATSTAT.

3) PATSTAT weekly file extracts from USPTO website:
Published Grants from November 22,nd 2005 until today;
Published Applications from September 29,th 2005 to today inclusive.
This data contains the sequence numbers.

4) Inventor & Applicant names for USPTO Published Applications from March 1,st 2001 to September 22,nd 2005 from DOCDB, data-format="docdba". This data contains the sequence numbers.

5) all other names from DOCDB, data-format="docdba". This data contains the sequence numbers.

Source field name

```
<inventor sequence="1" data-format="docdba">
      <inventor-name>
            <name>WHITTUM- HUDSON, JUDITH A</name>
      </inventor-name>
</inventor>
<inventor sequence="2" data-format="docdba">
      <inventor-name>
            <name>MACDONALD,ALEX BRUCE,</name>
      </inventor-name>
</inventor>
<inventor sequence="3" data-format="docdba">
      <inventor-name>
            <name>AN, LING LING</name>
      </inventor-name>
</inventor>
```

Source sub-field identifier

```
data-format="docdba"
```

Comments
An entry with a value 1 to n represents an inventor, an entry with the value 0 does not represent an inventor, but another person (e. g. an applicant).It is possible that there are applications with no inventors.
Consequently, adding the condition "INVT_SEQ_NR > 0" to the WHERE clause in a query retrieves only those persons from TLS207_PERS_APPLN or TLS227_PERS_PUBLN which are inventors.

Likewise, adding the condition "APPLT_SEQ_NR > 0 AND INVT_SEQ_NR > 0 " retrieves only persons which for a certain application are applicants as well as inventors.

For US data : Documents published after 1976-01-01: For the inventors , the sequence numbers are all given arbitrarily, with the exception of the documents published after March 2005, where the sequence numbers are all correct.

For all US documents published before 1976-01-01, where the data was taken from DOCDB, the sequence numbers are believed to be correct

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 19-04-2005 - First version
R. Heijna - 07-07-2005 - Value zero for the physical model
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-10-2013 - Changed source from EPO Bulletin to EP Register; changed domain

6.83 IPC

Name: IPC subclass / IPC main group
Also Known As: n/a
Description: First 4 – 8 characters of an IPC symbol according to WIPO ST.3. (In most cases they are only the first 4 characters)
Domain: Up to 8 ASCII characters; Example: `'B01D','A61K  6'`
Default value: n/a
Source database: See Eurostat's paper described in section 5.28 "TLS902_IPC_NACE2: Mapping between IPC and industrial sectors".

Source field name: n/a
Source sub-field identifier: n/a
Comments: See attribute IPC_CLASS_SYMBOL for the full IPC symbol.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2015 – Overall amendment, because the attribute has been moved from table INDUSTRY_IPC to table TLS902_IPC_NACE.

6.84 IPC_CLASS_LEVEL

Name: IPC classification level indicator
Also Known As: n/a
Description: Denotes whether an authority classified either in the full IPC, in main groups or in sub classes only.
Domain: 1 character:

A = classification in the full IPC	e.g. `'H04Q  7/32', 'C07K  14/00'`
C = classification in main groups only	e.g. `'H04H  1/00', 'A61K  31/00'`
S = classification in subclasses only	e. g. `'H04H', 'A61K'`

Default value: n/a
Source database: DOCDB
Source field name

```
<classifications-ipcr>
     <classification-ipcr sequence="1">
          <text>A43C  11/00        20060101CFI20070118BHUS        </text>
     </classification-ipcr>
     <classification-ipcr sequence="2">
          <text>A43C  11/00        20060101AFI20070118BHUS        </text>
     </classification-ipcr>
</classifications-ipcr>
```

Source sub-field identifier
Position 28 of the source-field

```
......12345678901234567890123456789012345678901234567890
<text>A43C  11/00        20060101CFI20070118BHUS        </text>
```

These text strings are all 50 bytes long. See WIPO ST.8.
Take byte 28 as the value of IPC_CLASS_LEVEL.

Source sub-field identifier
position 28
Comments
See the description of table TLS209_APPLN_IPC on how the IPC symbols, which are allocated in DOCDB to publications, are de-duplicated and assigned to applications in PATSTAT.

Modification history
Author of update - Date of update - Explanation of update
J. Rollinson - 27-08-2007 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 31-03-2011 - Roll up of Core symbols to Advanced
D. Lingua - 07-10-2011 - Value "S" (Symbol) has been eliminated in DOCDB
M. Kracker - 15-10-2014 - All levels A, C, S are available.

6.85 IPC_CLASS_SYMBOL

Name: IPC classification symbol (IPC 8th edition)
Also Known As: (IPC) class, (IPC) classification
Description: Classification symbol according to the International Patent Classification, eights edition (entered into force January 1, 2006)
Domain: Up to 15 characters (A-Z, 0-9, /, space) as allowed by IPC;
Examples: `A61K`
`H04Q  7/32`
`C07K 14/00`
`C07D 405/06`
`H01M2220/20`
Note that spaces may be required on position 5-7, because the slash "/" is always on the 9th position. For more details see the table below.
Default value: n/a
Source database: DOCDB
Source field name:

```
<classifications-ipcr>
      <classification-ipcr sequence="1">
            <text>A43C  11/00        20060101CFI20070118BHUS        </text>
      </classification-ipcr>
      <classification-ipcr sequence="2">
            <text>A43C  11/00        20060101AFI20070118BHUS        </text>
      </classification-ipcr>
</classifications-ipcr>
```

Source sub-field identifier
positions 1-15 of the source-field
Source codes

```
......12345678901234567890123456789012345678901234567890
<text>A43C  11/00        20060101CFI20070118BHUS        </text>
```

These text strings are all 50 bytes long. See WIPO ST.8:

For the recording of IPC symbols on machine-readable records a field of 50 positions should be allotted for each symbol, the 50 positions of the field to be used as follows:

symbol, the 50 positions of the field to be used as follows: *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
9	Separating character	/ (“Slash”)
10 to 15	Subgroup (left aligned)	00,…,999999, blank
16 to 19	For future use	4 blanks
20 to 27	Version indicator	YYYYMMDD date format
28	Classification level	C,A,S

29	First or later position of symbol	F,L
30	Classification value (invention or additional)	I,A
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 (ST.3)
43-50	For future use	8 blanks

Take the first 15 bytes as the value of IPC_CLASS_SYMBOL. For each symbol, be sure to take the corresponding values of IPC_GENER_AUTH, IPC_VERSION, IPC_POSITION, IPC_VALUE and IPC_CLASS_LEVEL from the same `classification-ipcr` element.

Comments
See the description of table TLS209_APPLN_IPC on how the IPC symbols, which are allocated in DOCDB to publications, are de-duplicated and assigned to applications in PATSTAT.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 19-04-2005 - First version
J. Rollinson - Aug 2007 - Addition of “Advanced" symbols
D. Lingua - 16-04-2009 - Amended text
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 19-02-2010 - "Source code" description modified
J. Rollinson - 01-04-2011 - Core no longer maintained by WIPO
M. Kracker - 15-10-2014 – Comment updated.
M. Kracker - 01-04-2016 – Examples showing the correct format have been added

6.86 IPC_GENER_AUTH

Name: IPC generating authority
Also Known As: n/a
Description: Patent office that generated the IPC classification of the application concerned
Domain: 2 ASCII characters (A-Z), according to WIPO ST.3

Default value: n/a
Source database: DOCDB
Source field name

```
<classifications-ipcr>
      <classification-ipcr sequence="1">
            <text>A43C  11/00        20060101CFI20070118BHUS        </text>
      </classification-ipcr>
      <classification-ipcr sequence="2">
            <text>A43C  11/00        20060101AFI20070118BHUS        </text>
      </classification-ipcr>
</classifications-ipcr>
```

Source sub-field identifier
Position 41-42: Generating office AA, ZZ (ST.3)
Comments
See WIPO ST.8.
See the description of table TLS209_APPLN_IPC on how the IPC symbols, which are allocated in DOCDB to publications, are de-duplicated and assigned to applications in PATSTAT.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 31-10-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 15-05-2013 - Added exception to Domain
M. Kracker - 15-10-2014 – Comment updated; removed exception to domain.

6.87 IPC_MAINGROUP_SYMBOL

Name: IPC subclass or IPC main group
Also Known As: n/a
Description: The subclass (i.e. first 4 characters) or main group (i.e. first 8 characters) of an IPC symbol according to WIPO ST.3
Domain: 4 or 8 ASCII characters;
Examples: `'A61K'`
`'A61K  8'`
`'A61K 133'`
Note: Spaces are relevant, as with all IPC or CPC symbols.
Default value: n/a
Source database:
http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/xls/ipc_technology.xls

Source field name:
Column IPC_CODE of the above-mentioned Excel file (without trailing %-sign)

Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version

6.88 IPC_POSITION

Name: First or later position of symbol
Also Known As: n/a
Description: Indicates the position of the class symbol in the sequence of classes that form the classification
Domain: 1 character: F=first, L=later. space =unidentified
Default value: space
Source database: DOCDB
Source field name:

```
<classifications-ipcr>
      <classification-ipcr sequence="1">
            <text>A43C  11/00        20060101CFI20070118BHUS        </text>
      </classification-ipcr>
      <classification-ipcr sequence="2">
            <text>A43C  11/00        20060101AFI20070118BHUS        </text>
      </classification-ipcr>
</classifications-ipcr>
```

If there is a space in `<classification-ipcr>` in position 29, then record a space in PATSTAT in IPC_POSITION.

Source sub-field identifier
Position 29: First or later position of symbol F, L
Comments
See WIPO ST.8 for an explanation.

For patent authorities (e. g. USPTO) where the law entails the concept of "first" class, the first class symbol in a list of class symbols is the main class. For other authorities, like the EPO, there is no meaning in the position - classes may be quoted in alphabetical order for instance. Some researchers use a weighting technique to analyse by IPC.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 19-04-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
J. Rollinson - 01-04-2011 - Core no longer maintained by WIPO
M. Kracker - 15-10-2014 – Comment updated

6.89 IPC_VALUE

Name: Classification value
Also Known As: Invention / Additional; Inventive/Non-Inventive
Description: Indication of the value of the classification i.e. is the class symbol relating to the invention or to aspects not related to the invention (but in the application).
Domain: 1 character: I=Invention, N=Additional (Non-Invention)
Default value: n/a
Source database: DOCDB
Source field name:

```
<classifications-ipcr>
      <classification-ipcr sequence="1">
            <text>A43C  11/00        20060101CFI20070118BHUS        </text>
      </classification-ipcr>
      <classification-ipcr sequence="2">
            <text>A43C  11/00        20060101AFI20070118BHUS        </text>
      </classification-ipcr>
</classifications-ipcr>
```

Source sub-field identifier
Position 30: Classification value (inventive or non-inventive) I, N
Comments
See WIPO ST.8 for an explanation.

See the description of table TLS209_APPLN_IPC on how the IPC symbols, which are allocated in DOCDB to publications, are de-duplicated and assigned to applications in PATSTAT.

Invention related IPC symbols are printed bold on the front page of patent documents, according to WIPO standard ST.10/C.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 19-04-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
J. Rollinson - 01-04-2011 - Core no longer maintained by WIPO
D. Lingua - 16-08-2012 - Remark on bold prints added
M. Kracker - 15-10-2014 – Comment updated

6.90 IPC_VERSION

Name: IPC version
Also Known As: n/a
Description: Version of the IPC
Domain: Date between '2006-01-01' and current date
Default value: n/a
Source database: DOCDB
Source field name:

```
<classifications-ipcr>
      <classification-ipcr sequence="1">
            <text>A43C  11/00        20060101CFI20070118BHUS        </text>
      </classification-ipcr>
      <classification-ipcr sequence="2">
            <text>A43C  11/00        20060101AFI20070118BHUS        </text>
      </classification-ipcr>
</classifications-ipcr>
```

Source sub-field identifier
Position 20 to 27: Version indicator in YYYYMMDD date format
Comments
See WIPO ST.8 for an explanation.

See the description of table TLS209_APPLN_IPC on how the IPC symbols, which are allocated in DOCDB to publications, are de-duplicated and assigned to applications in PATSTAT.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 19-04-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
J. Rollinson - 01-04-2011 - Core no longer maintained by WIPO
M. Kracker - 15-10-2014 - Comment updated

6.91 IPR_TYPE

Name: Type of Intellectual Property Right
Also Known As: n/a
Description: Type of Intellectual Property Right
Domain: 2 ASCII characters: PI, UM, DP;
PI - Patent of Invention
UM - Utility Model
DP - Design Patent
Default value: n/a
Source database: PATSTAT
Source field name: APPLN_AUTH, APPLN_KIND, PUBLN_KIND
Source sub-field identifier: n/a
Source codes
If first character of APPLN_KIND is ‘U’ or ‘V’ or ‘Y’ or ‘Z’, or
(APPLN_AUTH = 'FR' and APPL_KIND = 'A' and at least one related publication has a PUBLN_KIND = 'A3' or 'A4' or 'A7' or A8')
then IPR_TYPE = '**UM**' for utility model
else if APPLN_KIND = 'F ' and APPLN_AUTH is not 'FR' then IPR_TYPE = '**DP**' for design patent.
For all other values of APPLN_KIND, set IPR_TYPE to '**PI**' for Patent of Invention. Note that in America, a Patent of Invention is known as a Utility Patent.
This rule applies to all instances of APPLN_KIND, whether it is derived from application-reference or a priority-reference.

Comments
The rule to compute utility models and design patents does cover all major, but not necessarily all cases. The rule may be improved in the future.
Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 12-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML.
Added Design Patent info.
M. Kracker - 01-10-2013 - Changed rule to compute the IPR_TYPE; added comment

6.92 ISO_ALPHA3

Name: 3-letter country code
Also Known As: ISO 3166 alpha-3 code for countries/territories
Description: The three-letter code for the representation of countries/territories, as defined in standard ISO 3166
Domain: 3 ASCII letters [A-Z] (for countries/territories) or space (for intergovernmental organisations)
Default value: space
Source database:
ISO 3166 alpha-3 codes:
http://www.nationsonline.org/oneworld/country_code_list.htm
Deleted entries:
See section “Deleted Codes” in http://en.wikipedia.org/wiki/ISO_3166-1_alpha-3
Source field name
n/a
Source sub-field identifier
n/a
Comments

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2015 - First version

6.93 JP_CLASS_SCHEME

Name: Description of the JP classification scheme
Also Known As: n/a
Description:
The two schemes for JP classification are:
FI - File Index
FTERM - File Term
Domain: Up to 5 ASCII characters: FI, FTERM
Default value: n/a
Source database: DOCDB
Source field name

```
<patent-classifications>
    <patent-classification sequence="1">
        <classification-scheme office="JP" scheme="FI"/>
        <classification-symbol>4F21V8 /00 231</classification-symbol>
    </patent-classification>
    <patent-classification sequence="2">
        <classification-scheme office="JP" scheme="FTERM"/>
        <classification-symbol>4H129/BA20</classification-symbol>
    </patent-classification>
</patent-classifications>
```

Source sub-field identifier
n/a
Comments
These classifications are being stored in DOCDB as supplied by the National Office without inspection of the contents. The EPO does not hold any responsibility for content, format or validity.
Modification history
Author of update - Date of update - Explanation of update
D. Lingua - 04-08-2011 - First version

6.94 JP_CLASS_SYMBOL

Name: Symbols defined within the JP classification scheme
Also Known As: n/a
Description: The two schemes FI and FTERM consist of symbols, which can be up to 50 characters long.
Domain: Up to 50 characters (almost all symbols are between 10 and 18 characters long)
Default value: n/a
Source database: DOCDB
Source field name

```
<patent-classifications>
    <patent-classification sequence="1">
        <classification-scheme office="JP" scheme="FI"/>
        <classification-symbol>4F21V8 /00 231</classification-symbol>
    </patent-classification>
    <patent-classification sequence="2">
        <classification-scheme office="JP" scheme="FTERM"/>
        <classification-symbol>4H129/BA20</classification-symbol>
    </patent-classification>
</patent-classifications>
```

For JP_CLASS_SCHEME = FI, the first character (which is in the source data always “4”) does not have any significance and is removed.

For JP_CLASS_SCHEME = FTERM, the first character (a digit between 2 and 5) gives the broad technical area:
2 Residual technology
3 Mechanics
4 Chemistry
5 Electricity

Source sub-field identifier
n/a
Comments
These classifications are being stored in DOCDB as supplied by the National Office without inspection of the contents. The EPO does not hold any responsibility for content, format or validity.
Modification history
Author of update - Date of update - Explanation of update
D. Lingua - 04-08-2011 - First version
M. Kracker - 01-04-2021 – Removed first character in FI symbols;
Explained first character in FTERM symbols

6.95 LAPSE_COUNTRY

Name: Lapsed country or territory
Also Known As: n/a
Description: Office where the granted EP application has lapsed.
Domain: 2 ASCII characters (A-Z), according to WIPO ST.3
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/notification-of-lapse/@country

```
<legal-event providing-office="EP" date-added="20120802" date-previous-
exchange="20120802" sequence-number="12">
     <event-date>20120731</event-date>
     <event-code>PG25</event-code>
     <event-details>
          <event-description event-description-type="original">LAPSED IN A
CONTRACTING STATE ANNOUNCED VIA POSTGRANT INFORM. FROM NAT. OFFICE TO EPO</event-
description>
          <event-description lang="en">LAPSED IN A CONTRACTING STATE ANNOUNCED VIA
POSTGRANT INFORM. FROM NAT. OFFICE TO EPO</event-description>
          <notification-of-lapse country="HR">
               <date-patent-lapsed>20120621</date-patent-lapsed>
               <text>LAPSE BECAUSE OF FAILURE TO SUBMIT A TRANSLATION OF THE
DESCRIPTION OR TO PAY THE FEE WITHIN THE PRESCRIBED TIME-LIMIT</text>
          </notification-of-lapse>
     </event-details>
</legal-event>
```

Comments
This attribute is populated if and only if EVENT_CODE has the value PG25, VS25 or PG2D,

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version
M. Kracker - 01-04-2021 – Will also be populated for EVENT_CODE VS25 and PG2D

6.96 LAPSE_DATE

Name: Date of lapse
Also Known As: n/a
Description: Date when the lapse of a patent became effective.
Domain: Date
Default value: 9999-12-31

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/notification-of-lapse/date-patent-lapsed

```
<legal-event providing-office="EP" date-added="20120802" date-previous-
exchange="20120802" sequence-number="12">
     <event-date>20120731</event-date>
     <event-code>PG25</event-code>
     <event-details>
          <event-description event-description-type="original">LAPSED IN A
CONTRACTING STATE ANNOUNCED VIA POSTGRANT INFORM. FROM NAT. OFFICE TO EPO</event-
description>
          <event-description lang="en">LAPSED IN A CONTRACTING STATE ANNOUNCED VIA
POSTGRANT INFORM. FROM NAT. OFFICE TO EPO</event-description>
          <notification-of-lapse country="NO">
               <date-patent-lapsed>20120621</date-patent-lapsed>
               <text>LAPSE BECAUSE OF FAILURE TO SUBMIT A TRANSLATION OF THE
DESCRIPTION OR TO PAY THE FEE WITHIN THE PRESCRIBED TIME-LIMIT</text>
          </notification-of-lapse>
     </event-details>
</legal-event>
```

Comments
This attribute is populated if and only if EVENT_CODE has the value PG25 or VS25.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version
M. Kracker - 01-04-2021 – Will also be populated for EVENT_CODE VS25

6.97 LAPSE_TEXT

Name: Additional information about a lapse
Also Known As: n/a
Description: Additional information in free form text about the lapse of a patent.
Domain: up to 1 000 characters
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/notification-of-lapse/text

```
<legal-event providing-office="EP" date-added="20120503" date-previous-
exchange="20120503" sequence-number="19">
     <event-date>20120430</event-date>
     <event-code>PG25</event-code>
     <event-details>
          <event-description event-description-type="original">LAPSED IN A
CONTRACTING STATE ANNOUNCED VIA POSTGRANT INFORM. FROM NAT. OFFICE TO EPO</event-
description>
          <event-description lang="en">LAPSED IN A CONTRACTING STATE ANNOUNCED VIA
POSTGRANT INFORM. FROM NAT. OFFICE TO EPO</event-description>
          <notification-of-lapse country="IS">
               <date-patent-lapsed>20120302</date-patent-lapsed>
               <text>LAPSE BECAUSE OF FAILURE TO SUBMIT A TRANSLATION OF THE
DESCRIPTION OR TO PAY THE FEE WITHIN THE PRESCRIBED TIME-LIMIT</text>
          </notification-of-lapse>
     </event-details>
</legal-event>
```

Comments
This attribute is only populated if EVENT_CODE has the value PG25 or VS25.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version
M. Kracker - 01-04-2021 – Can also be populated for EVENT_CODE VS25

6.98 LAST_NAME

Name: Last name / Organization name
Also Known As: n/a
Description: Contains the last name (family name, surname) of a physical person or the name of a legal person
Domain: Up to 500 characters
Default value: empty string
Source database: USPTO data of published applications and published grants
Source field name:
`<addressbook> <last-name>`; if empty then `<addressbook> <orgname>`
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2016 – Domain extended to 500 characters

6.99 MIDDLE_NAME

Name: Middle name of a physical person
Also Known As: n/a
Description: Contains the middle name of a physical person
Domain: Up to 500 characters
Default value: empty string
Source database: USPTO data of published applications and published grants
Source field name:
`<addressbook> <middle-name>`
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2016 – Domain extended to 500 characters

6.100 NACE2_CODE

Name: 2-4-digit code of the Statistical Classification of Economic Activities in the European Community (Nomenclature statistique des activités économiques dans la CE)
Also Known As: n/a
Description: The 2-4 digits NACE2 code, like `'17'`, `'18.1'`, `'20.51'` or `'20.60'`.
Domain: Up to 5 ASCII characters;
It must not be defined as numerical field because trailing zeros are significant (e. g. NACE2 codes "`20.6`" and "`20.60`" are not the same).
Default value: n/a
Source database: See Eurostat's paper described in section 5.28 "TLS902_IPC_NACE2: Mapping between IPC and industrial sectors".

Source field name: n/a
Source sub-field identifier: n/a
Comments: This is a classification according to *industries*. A classification according to *technology* is the TECHN_FIELD_NR which can be found in the tables TLS901_TECHN_FIELD_IPC and TLS209_APPLN_IPC.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2015 - First version

6.101 NACE2_DESCR

Name: Description of the NACE2 code
Also Known As: n/a
Description: Short description of the NACE2 code

Domain: Up to 150 ASCII characters
Default value: n/a
Source database: See Eurostat's paper cited in section 5.28 "TLS902_IPC_NACE2: Mapping between IPC and industrial sectors".

Source field name: n/a
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2015 - First version

6.102 NACE2_WEIGHT

Name: Indicator whether a NACE2 code will be assigned to a certain IPC
Also Known As: n/a
Description: Weight (number 1 or 0) indicating whether there is a mapping between a particular IPC and a NACE2 code.

Domain: Number 0 or 1
Default value: 1
Source database: See Eurostat's paper cited in section 5.28 "TLS902_IPC_NACE2: Mapping between IPC and industrial sectors".

Source field name: n/a
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2015 - First version
M. Kracker - 01-04-2016 – Changed domain and description

6.103 NAME_FREEFORM

Name: Full name in a single string
Also Known As: n/a
Description: Contains the full name in case the name is not available in structured form, where first, middle and last name are in different data fields.
Domain: Up to 500 characters
Default value: empty string
Source database: DOCDB and EP Register

For DOCDB data:
Source field name

```
<inventors>
      <inventor sequence="1" data-format="docdb">
            <inventor-name>
                  <name>MACDONALD ALEX BRUCE</name>
            </inventor-name>
            <residence>
                  <country>US</country>
            </residence>
      </inventor>
      <inventor sequence="2" data-format="docdb">
            <inventor-name>
                  <name>AN LING LING</name>
            </inventor-name>
            <residence>
                  <country>US</country>
            </residence>
      </inventor>
      <inventor sequence="1" data-format="docdba">
            <inventor-name>
                  <name>WHITTUM- HUDSON, JUDITH A</name>
            </inventor-name>
      </inventor>
      <inventor sequence="2" data-format="docdba">
            <inventor-name>
                  <name>MACDONALD,ALEX BRUCE,</name>
            </inventor-name>
      </inventor>
      <inventor sequence="3" data-format="docdba">
            <inventor-name>
                  <name>AN, LING LING</name>
            </inventor-name>
      </inventor>
</inventors>

<applicants>
      <applicant sequence="1" data-format="docdba">
            <applicant-name>
                  <name>THE JOHNS HOPKINS UNIVERSITY MACDONALD, ALEX BRUCE</name>
            </applicant-name>
      </applicant>
      <applicant sequence="2" data-format="docdba">
            <applicant-name>
                  <name>AN, LING LING UNIVERSITY OF MASSACHUSETTS, A
PUBLIC INSTITUTION OF HIGHER EDUCATION OF THE COMMONWEALTH
OF MASSACHUSETTS,</name>
            </applicant-name>
      </applicant>
```

```
</applicants>
```

Example where only an original name exists:

```
<exch:applicants>
      <exch:applicant sequence="1" data-format="original">
            <exch:applicant-name>
                  <name>디디에르-베르케 아게</name>
            </exch:applicant-name>
      </exch:applicant>
</exch:applicants>
```

Source sub-field identifier
It occurs that DOCDB contains the names in DOCDB standardised format, but not in unstandardised format. So the first applicable rule of the following ordered rules must be executed:

1. If **un**standardised applicant / inventor name exists, take format "docdb**a**".
2. If standardised applicant / inventor name exists, take format "docdb".
3. If original applicant / inventor name exists, take format "original".

For EP Register data:
Source field name

```
<parties>
      <applicants change-gazette-num="2000/29">
            <applicant app-type="applicant" designation="all" sequence="1">
                  <addressbook>
                        <name>Seidel, Helmut</name>
                        <address>
                              <address-1>Fliederstrasse 19</address-1>
                              <address-2>65396 Walluf</address-2>
                              <country>DE</country>
                        </address>
                  </addressbook>
                  <nationality>
                        <country/>
                  </nationality>
                  <residence>
                        <country/>
                  </residence>
            </applicant>
      </applicants>
      <inventors change-gazette-num="2000/29">
            <inventor sequence="01">
                  <addressbook>
                        <name>Franta, Georg</name>
                        <address>
                              <address-1>Ulrich-Rapp-Strasse 18</address-1>
                              <address-2>87634 Obergünzburg</address-2>
                              <country>DE</country>
                        </address>
                  </addressbook>
            </inventor>
            <inventor sequence="02">
                  <addressbook>
                        <name>Dojan, Viktor</name>
```

```
                        <address>
                              <address-1>Ludwig-Strecker-Strasse 5</address-1>
                              <address-2>55129 Mainz</address-2>
                              <country>DE</country>
                        </address>
                  </addressbook>
            </inventor>
      </inventors>
```

Source sub-field identifier: n/a

Comments:
DOCDB data in `data-format = "docdba"` are stored in PATSTAT "as received" by the EPO from other offices after converting to upper case and removing diacritics.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2014 – DOCDB data: Source sub-field identifier changed.

6.104 NAT_CLASS_SYMBOL

Name: National classification symbol
Also Known As: n/a
Description: Classification symbol according to a National classification scheme
Domain: Up to 15 characters
Default value: n/a
Source database: DOCDB
Source field name:

```
<classification-national>
      <text>002002500</text>
      <text>X002410000</text>
</classification-national>
```

If a string of symbols contains a comma, then split the string at the comma and create multiple entries.

Source sub-field identifier
n/a
Comments
These symbols are stored in PATSTAT against the APPLN_ID.

National classification is found in DOCDB mainly for AT, BR, CA, CH, DE, DK, GB and MX. JP and US national classification symbols are in tables TLS222_APPLN_JP_CLASS and TLS223_APPLN_DOCUS.

These national classification symbols are stored exactly as received by the EPO. No corrections are made.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 01-07-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 19-02-2010 - Modified comments
D. Lingua - 11-10-2011 - US and JP symbols have been moved to separate tables

6.105 NAT_PHASE

Name: Indicator whether the application is in the national phase
Also Known As: n/a
Description: Indicates that an application is in the national phase.
Domain: 1 ASCII character

Y	Yes
N	No
space	not known (In case of uncertain interpretations; used very little or not at all)

Default value: N

Source database: PATSTAT
Source field name: Derived from table TLS201_APPLN

Y	if the application has APPLN_KIND <> W and APPLN_AUTH is a national office;
N	otherwise

Source sub-field identifier: n/a

Comments
For explanation and disclaimer see attribute INT_PHASE in section 6.80.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2016 - First version

6.106 NB_APPLICANTS

Name: Number of applicants of an application
Also Known As: n/a
Description: Number of applicants of an application according to the most recent publication which contains Latin person names
Domain: Number 0 … about 250
Default value: n/a
Source database: PATSTAT
Source field name: Derived from table TLS207_PERS_APPLN

Source sub-field identifier: n/a

Comments: If no publication of the application contains applicant names in Latin characters, then NB_APPLICANTS will be zero.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-10-2019 - Change in data content: recent publication must contain Latin applicant names

6.107 NB_CITING_DOCDB_FAM

Name: Number of forward citations on family level
Also Known As: n/a
Description: Number of distinct DOCDB simple families citing at least one of the publications or applications of the DOCDB simple family of the current application (search report citations from TLS212_CITATION)
Domain: Number 0 .. about 3.000
Default value: n/a
Source database: PATSTAT
Source field name: Derived from table TLS228_DOCDB_FAM_CITN

Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-04-2015 - Name of attribute changed for clarification (was: NB_CITATION)

6.108 NB_INVENTORS

Name: Number of inventors of an application
Also Known As: n/a
Description: Number of inventors of an application according to the most recent publication which contains Latin person names
Domain: Number 0 … about 250
Default value: n/a
Source database: PATSTAT
Source field name: Derived from table TLS207_PERS_APPLN

Source sub-field identifier: n/a

Comments: If no publication of the application contains inventor names in Latin characters, then NB_INVENTORS will be zero.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-10-2019 - Change in data content: recent publication must contain Latin inventor names

6.109 NOT_WITH_IPC

Name: IPC main group not co-occurring with IPC sub class
Also Known As: n/a
Description: Empty or first 8 characters of an IPC symbol according to WIPO ST.8.

Domain: Up to 8 ASCII characters; Example: `'A61K  6'`
Default value: empty
Source database: See Eurostat's paper described in section 5.28 "TLS902_IPC_NACE2: Mapping between IPC and industrial sectors".

Source field name: n/a
Source sub-field identifier: n/a
Comments: IPC main group which must not co-occur with the IPC in attribute IPC.
In the most cases this field is empty

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2015 - First version

6.110 NPL_ABSTRACT_NR

Name: Identifier for the NPL abstract
Also Known As: n/a
Description: Identifier, e.g. for Chemical Abstracts (CAS), Patent Abstract of Japan (PAJ), citation of a periodical publication, ...
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Chemical Abstract citation (c)
- Biological abstract citation (i)
- Patent Abstracts of Japan (j)
- Derwent citation (d)
- Database citation (e)
- World Wide Web / Internet search citation (w)

Domain: up to 50 ASCII characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = c, i, j
/nplcit/article/absno

```
<nplcit num="1" npl-type="c">
     <text>CHEMICAL ABSTRACTS, Columbus, Ohio, US; abstract no. 77-92-2</text>
     <article>
          <serial>
               <sertitle/>
          </serial>
          <absno>77-92-2</absno>
     </article>
</nplcit>
```

2) For NPL_TYPE = d, e, w
/nplcit/online/absno

```
<nplcit num="1" npl-type="d" extracted-xp="002556425">
     <text>DATABASE WPI Week 200235, Derwent Publications Ltd., London, GB; AN
2002-309048, XP002556425</text>
     <online>
          <edition>0</edition>
          <vid>2002</vid>
          <ino>35</ino>
          <absno>2002-309048</absno>
     </online>
     <source-doc>
          <document-id>
               <country>JP</country>
               <doc-number>2001288238</doc-number>
               <kind>A</kind>
          </document-id>
     </source-doc>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.111 NPL_AUTHOR

Name: Author
Also Known As: n/a
Description: Name of the author.
Special meaning in these cases:
- For Derwent citations (NPL_TYPE = d) and Database citations (NPL_TYPE = e): Where there is a Patent citation embedded, this attribute contains the name of an applicant or inventor; otherwise it is the name of the author of the article or abstract.
- For WWW / Internet search citations (NPL_TYPE = w): The (person) name of such article is used as "author name".

This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Chemical Abstract citation (c)
- Biological abstract citation (i)
- Serial / Journal / Periodical citation (s)
- Derwent citation (d)
- Database citation (e)
- World Wide Web / Internet search citation (w)

Domain: up to 1 000 characters.
Multiple authors are typically indicated by "ET AL" or separated by a semicolon
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b, c, i, s
/nplcit/article/author/name

```
<nplcit num="20" npl-type="b">
     <text>MURRAY: "Methods in Molecular Biology, Volume 7: Gene Transfer and
Expression Protocols", vol. 7, 1991, THE HUMANA PRESS, INC., article BAILEY ET
AL.: "Manipulation of Baculovirus Vectors", pages: 147 - 168</text>
     <article>
          <author>
               <name>BAILEY ET AL.</name>
          </author>
          <atl>Manipulation of Baculovirus Vectors</atl>
          <book>
               <author>
                    <name>MURRAY</name>
               </author>
               <book-title>Methods in Molecular Biology, Volume 7: Gene Transfer
and Expression Protocols</book-title>
               <imprint>
                    <name>THE HUMANA PRESS, INC.</name>
                    <pubdate>1991</pubdate>
               </imprint>
               <vid>7</vid>
               <location>
                    <pp>
                         <ppf>147</ppf>
                         <ppl>168</ppl>
```

```
                    </pp>
               </location>
          </book>
     </article>
</nplcit>
```

2) For NPL_TYPE = d, e, w
/nplcit/online/author/name

```
<nplcit num="1" npl-type="d" extracted-xp="002551011">
     <text>DATABASE WPI Week 200818, Derwent Publications Ltd., London, GB; AN
2008-C41297, XP002551011, OBIKAWA T: "Organic electroluminescent ... crystalline
mesogen"</text>
     <online>
          <author>
               <name>OBIKAWA T</name>
          </author>
          <online-title>Organic electroluminescent ... crystalline mesogen
</online-title>
          <edition>0</edition>
          <vid>2008</vid>
          <ino>18</ino>
          <absno>2008-C41297</absno>
     </online>
     <source-doc>
          <document-id>
               <country>JP</country>
               <doc-number>2005222777</doc-number>
               <kind>A</kind>
          </document-id>
     </source-doc>
</nplcit>
```

Comments
For books (NPL_TYPE = b): see also attribute NPL_EDITOR.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version
M. Kracker - 01-10-2017 – Extend domain to 1 000 characters

6.112 NPL_BIBLIO

Name: Non-Patent Literature bibliography
Also Known As: n/a
Description: Bibliographic data of the Non-Patent Literature
Domain: Up to 4 000 characters. In the PATSTAT 2018 Spring Edition the longest string was 3 800 characters
Default value: n/a
Source database: DOCDB
Source field name:
Each NPL citation leads to a record in the PATSTAT Non-Patent Literature table. Ignore any text which is empty (spaces), and in these cases take care with the calculation of the NPL_CITN_SEQ_NR and the CITN_ID.

Example from publication EP 1944010 A2:

```
<exch:references-cited>
...
      <exch:citation cited-phase="APP" cited-date="20080116" sequence="2">
            <nplcit num="1" npl-type="b">
                  <text>KH. SCHRADER: &quot;Grundlagen und Rezepturen der
Kosmetika&quot;, vol. 2, 1989, DR. ALFRED H&#220;THIG VERLAG</text>
                  <article>
                        <book>
                              <author>
                                    <name>KH. SCHRADER</name>
                              </author>
                              <book-title>Grundlagen und Rezepturen der
Kosmetika</book-title>
                              <imprint>
                                    <name>DR. ALFRED H&#252;THIG VERLAG</name>
                                    <pubdate>1989</pubdate>
                              </imprint>
                              <vid>2</vid>
                        </book>
                  </article>
            </nplcit>
      </exch:citation>
      <exch:citation cited-phase="APP" cited-date="20080116" sequence="3">
            <nplcit num="2" npl-type="b">
                  <text>W. UMBACH: &quot;Kosmetik&quot;, vol. 2, 1995, GEORG
THIEME VERLAG</text>
                  <article>
                        <book>
                              <author>
                                    <name>W. UMBACH</name>
                              </author>
                              <book-title>Kosmetik</book-title>
                              <imprint>
                                    <name>GEORG THIEME VERLAG</name>
                                    <pubdate>1995</pubdate>
                              </imprint>
                              <vid>2</vid>
                        </book>
                  </article>
            </nplcit>
      </exch:citation>
...
```

```
</exch:references-cited>
```

Source sub-field identifier
n/a
Comments
For NPL citations (e.g. Derwent abstracts, Patent Abstracts of Japan) which include a reference to a patent document see description of table TLS212_CITATION and the source field name in attribute description of CITED_PAT_PUBLN_ID.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 01-07-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-04-2015 – Added “Corresponding documents” to comment
M. Kracker - 01-04-2018 – Removed “Corresponding documents” from comment
M. Kracker - 01-10-2018 – Domain extended to allow for longer strings

6.113 NPL_CITN_SEQ_NR

Name: Sequence number of the NPL citation
Also Known As: n/a
Description: Number for an NPL (non-patent literature) citation in the series of NPL citations for one publication/origin combination
Domain: Number 0 … about 1000
Default value: 0
Source database: Computed from PATSTAT. It is a sequential number for each NPL citation. The numbering starts with 1 for each origin of citations (CITN_ORIGIN).
The NPL_CITN_SEQ_NR will be set to 0 when the citation is not an NPL citation, but a patent citation.

Source field name: n/a

Source sub-field identifier: n/a

Comments
The NPL_CITN_SEQ_NR attribute does ***not*** indicate the order of appearance of NPL citations.

The sequence numbers start at 1 for each origin of the citations.

See also attributes PAT_CITN_SEQ_NR and CITN_ID.

References in the EP Search Report to the original WO search report publication are also included as NPL citations, because most likely the EP search report did not repeat the references which were cited in the original PCT search report. E.g. see references of EP1468879A1, which contains an NPL citation with the text "See also references of WO 03064220A1".

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 09-05-2005 - First version CITN_EXTRACT-NPL
R. Heijna - 21-11-2005 - CITN_EXTRACT-NPL deleted
R. Heijna - 21-11-2005 - First version NPL_CITN_SEQ_NR
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-12-2015 - Changed processing instructions and comments

6.114 NPL_DOI

Name: Digital Object Identifier
Also Known As: DOI
Description: A persistent identifier used to uniquely identify electronic documents, e.g. journal articles (see https://www.doi.org/).
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Serial / Journal / Periodical citation (s)
- World Wide Web / Internet search citation (w)
Domain: up to 500 ASCII characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b
/nplcit/article/book/doi

Example from EP 2164226 A1:

```
<nplcit num="2" npl-type="b" extracted-xp="010860584">
     <text>B. CARMINATI; E. FERRARI; P. C. K. HUNG.: &quot;In WIRI&apos;05:
Proceedings of the International Workshop on Challenges in Web Information
Retrieval and Integration&quot;, 2005, IEEE COMPUTER SOCIETY, article &quot;Web
service composition: A security perspective&quot;, pages: 248 - 253, XP010860584,
DOI: doi:10.1109/WIRI.2005.36</text>
     <article>
          <atl>Web service composition: A security perspective</atl>
          <book>
               <author>
                    <name>B. CARMINATI; E. FERRARI; P. C. K. HUNG.</name>
               </author>
               <book-title>In WIRI'05: Proceedings of the International Workshop
on Challenges in Web Information Retrieval and Integration</book-title>
               <imprint>
                    <name>IEEE COMPUTER SOCIETY</name>
                    <pubdate>2005</pubdate>
               </imprint>
               <location>
                    <pp>
                         <ppf>248</ppf>
                         <ppl>253</ppl>
                    </pp>
               </location>
               <doi>doi:10.1109/WIRI.2005.36</doi>
          </book>
     </article>
</nplcit>
```

2) For NPL_TYPE = s
/nplcit/article/serial/doi

```
<nplcit num="20" npl-type="s" extracted-xp="008150479">
     <text>BERNIER UR; BOOTH MM; YOST RA.: "Analysis of human skin emanations by
gas chromatography/ mass spectrometry. 1.Thermal desorption of attractants for
the yellow fever mosquito (Aedes aegypti) from handled glass beads", ANALYTICAL
```

```
CHEMISTRY, vol. 71, 1999, pages 1 - 7, XP008150479, DOI:
doi:10.1021/ac980990v</text>
     <article>
          <author>
               <name>BERNIER UR; BOOTH MM; YOST RA.</name>
          </author>
          <atl>Analysis of human skin emanations by gas chromatography/ mass
spectrometry. 1.Thermal desorption of attractants for the yellow fever mosquito
(Aedes aegypti) from handled glass beads</atl>
          <serial>
               <sertitle>ANALYTICAL CHEMISTRY</sertitle>
               <pubdate>
                    <sdate>1999</sdate>
               </pubdate>
               <vid>71</vid>
               <doi>doi:10.1021/ac980990v</doi>
          </serial>
          <location>
               <pp>
                    <ppf>1</ppf>
                    <ppl>7</ppl>
               </pp>
          </location>
     </article>
</nplcit>
```

3) For NPL_TYPE = w
/nplcit/online/serial/doi

Example from EP 2152028 A1:

```
<nplcit num="1" npl-type="w" extracted-xp="032737684">
     <text>NINGNING HU; PETER STEENKISTE: &quot;Estimating Available Bandwidth
Using Packet Pair Probing&quot;, CMU-CS-02-166, 9 September 2002 (2002-09-09),
XP032737684, Retrieved from the Internet &lt;URL:www.cs.cmu.edu/-hnn/papers/igi-
tr.pdf.&gt; DOI: doi:10.1109/SMARTCOMP-W.2014.7046666</text>
     <online>
          <author>
               <name>NINGNING HU; PETER STEENKISTE</name>
          </author>
          <online-title>Estimating Available Bandwidth Using Packet Pair
Probing</online-title>
          <serial>
               <sertitle>CMU-CS-02-166</sertitle>
               <doi>doi:10.1109/SMARTCOMP-W.2014.7046666</doi>
          </serial>
          <pubdate>
               <sdate>20020909</sdate>
          </pubdate>
          <avail>www.cs.cmu.edu/-hnn/papers/igi-tr.pdf.</avail>
     </online>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.115 NPL_EDITOR

Name: Editor
Also Known As: n/a
Description: Name of the editor or the author of a book.
This attribute may only be populated for this NPL type (see attribute NPL_TYPE):
- Book citation (b)
Domain: up to 500 characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b
/nplcit/article/book/author/name

```
<nplcit num="20" npl-type="b">
     <text>MURRAY: "Methods in Molecular Biology, Volume 7: Gene Transfer and
Expression Protocols", vol. 7, 1991, THE HUMANA PRESS, INC., article BAILEY ET
AL.: "Manipulation of Baculovirus Vectors", pages: 147 - 168</text>
     <article>
          <author>
               <name>BAILEY ET AL.</name>
          </author>
          <atl>Manipulation of Baculovirus Vectors</atl>
          <book>
               <author>
                    <name>MURRAY</name>
               </author>
               <book-title>Methods in Molecular Biology, Volume 7: Gene Transfer
and Expression Protocols</book-title>
               <imprint>
                    <name>THE HUMANA PRESS, INC.</name>
                    <pubdate>1991</pubdate>
               </imprint>
               <vid>7</vid>
               <location>
                    <pp>
                         <ppf>147</ppf>
                         <ppl>168</ppl>
                    </pp>
               </location>
          </book>
     </article>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.116 NPL_ISBN

Name: International Standard Book Number (ISBN)
Also Known As: n/a
Description: International Standard Book Number (ISBN).
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Patent Abstracts of Japan (j)
- Serial / Journal / Periodical citation (s)
- World Wide Web / Internet search citation (w)
Domain: up to 30 ASCII characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b
/nplcit/article/book/isbn

```
<nplcit num="1" npl-type="b" extracted-xp="010779133">
     <text>RAHMAN M ET AL: "Medical Image Retrieval and Registration: Towards
Computer Assisted Diagnostic Approach", 1 September 2004, MEDICAL INFORMATION
SYSTEMS: THE DIGITAL HOSPITAL, 2004. IDEAS '04-DH. PROCEEDINGS. IDEAS WORKSHOP ON
BEIJING, CHINA 01-03 SEPT. 2004, PISCATAWAY, NJ, USA,IEEE, PAGE(S) 78 - 89, ISBN:
978-0-7695-2289-0, XP010779133</text>
     <article>
          <book>
               <author>
                    <name>RAHMAN M ET AL</name>
               </author>
               <book-title>Medical Image Retrieval and Registration: Towards
Computer Assisted Diagnostic Approach</book-title>
               <imprint>
                    <name>MEDICAL INFORMATION SYSTEMS: THE DIGITAL HOSPITAL, 2004.
IDEAS '04-DH. PROCEEDINGS. IDEAS WORKSHOP ON BEIJING, CHINA 01-03 SEPT. 2004,
PISCATAWAY, NJ, USA,IEEE, PAGE(S) 78 - 89</name>
                    <pubdate>20040901</pubdate>
               </imprint>
               <isbn>978-0-7695-2289-0</isbn>
          </book>
     </article>
</nplcit>
```

2) For NPL_TYPE = j, s
/nplcit/article/serial/isbn

```
<nplcit num="1" npl-type="s" extracted-xp="031221242">
     <text>SENFT C. ET AL: "Cross Sensitivity and Stability of FET - Based
Hydrogen Sensors", SENSORS, 2007 IEEE, IEEE, PI, 28 October 2007 (2007-10-28),
pages 1036 - 1039, XP031221242, ISBN: 978-1-4244-1261-7</text>
     <article>
          <author>
               <name>SENFT C. ET AL</name>
          </author>
          <atl>Cross Sensitivity and Stability of FET - Based Hydrogen
Sensors</atl>
          <serial>
```

```
            <sertitle>SENSORS, 2007 IEEE, IEEE, PI</sertitle>
            <pubdate>
                <sdate>20071028</sdate>
            </pubdate>
            <isbn>978-1-4244-1261-7</isbn>
        </serial>
        <location>
            <pp>
                <ppf>1036</ppf>
                <ppl>1039</ppl>
            </pp>
        </location>
    </article>
</nplcit>
```

3) For NPL_TYPE = w
/nplcit/online/serial/isbn

```
<nplcit num="1" npl-type="w" extracted-xp="002511547">
    <text>BECKER F., SCHERER A., WEIGOLD J., BRAUN M.: "a simple indirect
voltage sensing method for line-connected inverters", POWER ELECTRONICS AND
APPLICATIONS, 2007 EUROPEAN CONFERENCE ON, 2 September 2007 (2007-09-02) - 5
September 2007 (2007-09-05), pages 1 - 7, XP002511547, ISBN: 978-92-75815-10-8,
Retrieved from the Internet &lt;URL:ieeexplore.ieee.org&gt;</text>
    <online>
        <author>
            <name>BECKER F., SCHERER A., WEIGOLD J., BRAUN M.</name>
        </author>
        <online-title>a simple indirect voltage sensing method for line-
connected inverters</online-title>
        <serial>
            <sertitle>POWER ELECTRONICS AND APPLICATIONS, 2007 EUROPEAN
CONFERENCE ON</sertitle>
            <isbn>978-92-75815-10-8</isbn>
        </serial>
        <pubdate>
            <sdate>20070902</sdate>
            <edate>20070905</edate>
        </pubdate>
        <location>
            <pp>
                <ppf>1</ppf>
                <ppl>7</ppl>
            </pp>
        </location>
        <avail>ieeexplore.ieee.org</avail>
    </online>
</nplcit>
```

Comments
As of the 2017 Spring Edition, there are no values for this attribute for NPL_TYPE = j (Patent Abstracts of Japan)

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.117 NPL_ISSN

Name: International Standard Serial Number (ISSN)
Also Known As: n/a
Description: International Standard Serial Number (ISSN)
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Chemical Abstract citation (c)
- Biological abstract citation (i)
- Patent Abstracts of Japan (j)- Serial / Journal / Periodical citation (s)
- Derwent citation (d)
- World Wide Web / Internet search citation (w)

Domain: up to 30 ASCII characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b
/nplcit/article/book/issn

```
<nplcit num="1" npl-type="b" extracted-xp="002496773">
     <text>ROCHETTE ANNIE ET AL: "Genome-wide gene expression profiling analysis
of Leishmania major and Leishmania infantum developmental stages reveals
substantial differences between the two species.", 2008, BMC GENOMICS 2008, VOL.
9, PAGE(S) 255, ISSN: 1471-2164, XP002496773</text>
     <article>
          <book>
               <author>
                    <name>ROCHETTE ANNIE ET AL</name>
               </author>
               <book-title>Genome-wide gene expression profiling analysis of
Leishmania major and Leishmania infantum developmental stages reveals substantial
differences between the two species.</book-title>
               <imprint>
                    <name>BMC GENOMICS 2008, VOL. 9, PAGE(S) 255</name>
                    <pubdate>2008</pubdate>
               </imprint>
               <issn>1471-2164</issn>
          </book>
     </article>
</nplcit>
```

2) For NPL_TYPE = c, i, j, s
/nplcit/article/serial/issn

```
<nplcit num="1" npl-type="s" extracted-xp="004519684">
     <text>LV J ET AL: "Controlled growth of three morphological structures of
magnesium hydroxide nanoparticles by wet precipitation method", JOURNAL OF
CRYSTAL GROWTH, ELSEVIER, AMSTERDAM, NL, vol. 267, no. 3-4, 1 July 2004 (2004-07-
01), pages 676 - 684, XP004519684, ISSN: 0022-0248</text>
     <article>
          <author>
               <name>LV J ET AL</name>
          </author>
```

```
          <atl>Controlled growth of three morphological structures of magnesium
hydroxide nanoparticles by wet precipitation method</atl>
          <serial>
               <sertitle>JOURNAL OF CRYSTAL GROWTH, ELSEVIER, AMSTERDAM,
NL</sertitle>
               <pubdate>
                    <sdate>20040701</sdate>
               </pubdate>
               <vid>267</vid>
               <ino>3-4</ino>
               <issn>0022-0248</issn>
          </serial>
          <location>
               <pp>
                    <ppf>676</ppf>
                    <ppl>684</ppl>
               </pp>
          </location>
     </article>
</nplcit>
```

3) For NPL_TYPE = d
/nplcit/online/issn

Currently there is no example in the 2017 Spring Edition available.

4) For NPL_TYPE = w
/nplcit/online/serial/issn

```
<nplcit num="3" npl-type="w" extracted-xp="023059967">
     <text>NISHIYAMA Y ET AL: "Construction of a US3 lacZ insertion mutant of
herpes simplex virus type 2 and characterization of its phenotype in vitro and in
vivo", VIROLOGY, ACADEMIC PRESS,ORLANDO, US, vol. 190, no. 1, 1 September 1992
(1992-09-01), pages 256 - 268, XP023059967, ISSN: 0042-6822, [retrieved on
19920901]</text>
     <online>
          <author>
               <name>NISHIYAMA Y ET AL</name>
          </author>
          <online-title>Construction of a US3 lacZ insertion mutant of herpes
simplex virus type 2 and characterization of its phenotype in vitro and in
vivo</online-title>
          <serial>
               <sertitle>VIROLOGY, ACADEMIC PRESS,ORLANDO, US</sertitle>
               <vid>190</vid>
               <ino>1</ino>
               <issn>0042-6822</issn>
          </serial>
          <pubdate>
               <sdate>19920901</sdate>
          </pubdate>
          <location>
               <pp>
                    <ppf>256</ppf>
                    <ppl>268</ppl>
               </pp>
          </location>
          <srchdate>
               <date>19920901</date>
          </srchdate>
```

```
    </online>
</nplcit>
```

Comments
As of the 2017 Spring Edition, there are no values for this attribute for NPL_TYPE = i, j and d
Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.118 NPL_ISSUE

Name: Issue number
Also Known As: n/a
Description: Specifies the issue of a Non-Patent Literature.
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Chemical Abstract citation (c)
- Biological abstract citation (i)
- Patent Abstracts of Japan (j)
- Serial / Journal / Periodical citation (s)
- Derwent citation (d)
- World Wide Web / Internet search citation (w)

Domain: up to 50 ASCII characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = c, i, j, s
/nplcit/article/serial/ino

```
<nplcit num="1" npl-type="c" extracted-xp="002128422">
     <text>CHEMICAL ABSTRACTS, vol. 128, no. 9, 2 March 1998, Columbus, Ohio, US;
abstract no. 98937, BIESELER, BARBARA ET AL: "Maize selectivity of FOE 5043.
Degradation of active ingredient by glutathione-S-transferases"
XP002128422</text>
     <article>
          <author>
               <name>BIESELER, BARBARA ET AL</name>
          </author>
          <atl>Maize selectivity of FOE 5043. Degradation of active ingredient by
glutathione-S-transferases</atl>
          <serial>
               <sertitle/>
               <pubdate>19980302</pubdate>
               <vid>128</vid>
               <ino>9</ino>
          </serial>
          <absno>98937</absno>
     </article>
</nplcit>
```

2) For NPL_TYPE = d
/nplcit/online/ino

```
<nplcit num="1" npl-type="d" extracted-xp="002715933">
     <text>DATABASE WPI Week 200443, Derwent Publications Ltd., London, GB; AN
2004-460824, XP002715933</text>
     <online>
          <edition>0</edition>
          <vid>2004</vid>
          <ino>43</ino>
          <absno>2004-460824</absno>
     </online>
     <source-doc>
          <document-id>
```

```
            <country>WO</country>
            <doc-number>2004049053</doc-number>
            <kind>A1</kind>
        </document-id>
    </source-doc>
</nplcit>
```

3) For NPL_TYPE = w
/nplcit/online/serial/ino

```
<nplcit num="1" npl-type="w" extracted-xp="025543349">
                          <text>SAITO H ET AL: "Cytotoxicity of chlorophenols
to goldfish GFS cells with the MTT and LDH assays", TOXICOLOGY IN VITRO, ELSEVIER
SCIENCE, GB, vol. 8, no. 5, 1 October 1994 (1994-10-01), pages 1107 - 1112,
XP025543349, ISSN: 0887-2333, [retrieved on 19941001]</text>
    <online>
        <author>
            <name>SAITO H ET AL</name>
        </author>
        <online-title>Cytotoxicity of chlorophenols to goldfish GFS cells with
the MTT and LDH assays</online-title>
        <serial>
            <sertitle>TOXICOLOGY IN VITRO, ELSEVIER SCIENCE, GB</sertitle>
            <vid>8</vid>
            <ino>5</ino>
            <issn>0887-2333</issn>
        </serial>
        <pubdate>
            <sdate>19941001</sdate>
        </pubdate>
        <location>
            <pp>
                <ppf>1107</ppf>
                <ppl>1112</ppl>
            </pp>
        </location>
        <srchdate>
            <date>19941001</date>
        </srchdate>
    </online>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.119 NPL_PAGE_FIRST

Name: First page of citation
Also Known As: n/a
Description: Start of page range.
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Chemical Abstract citation (c)
- Biological abstract citation (i)
- Serial / Journal / Periodical citation (s)
- Derwent citation (d)
- World Wide Web / Internet search citation (w)
Domain: up to 200 ASCII characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b
/nplcit/article/book/location/pp/ppf

```
<nplcit num="3" npl-type="b">
     <text>"Bergey's manual of systematic Bacteriology.", vol. 2, 1986, WILLIAMS
&amp;amp; WILKINS, pages: 635</text>
     <article>
          <book>
               <book-title>Bergey's manual of systematic Bacteriology.</book-
title>
               <imprint>
                    <name>WILLIAMS &amp;amp; WILKINS</name>
                    <pubdate>1986</pubdate>
               </imprint>
               <vid>2</vid>
               <location>
                    <pp>
                         <ppf>635</ppf>
                    </pp>
               </location>
          </book>
     </article>
</nplcit>
```

2) For NPL_TYPE = c, i
/nplcit/article/location/pp

```
<nplcit num="1" npl-type="c" extracted-xp="002169245">
     <text>CHEMICAL ABSTRACTS, vol. 101, no. 25, 17 December 1984, Columbus,
Ohio, US; abstract no. 230466q, page 765; column R; XP002169245</text>
     <article>
          <serial>
               <sertitle/>
               <pubdate>19841217</pubdate>
               <vid>101</vid>
               <ino>25</ino>
          </serial>
          <absno>230466q</absno>
```

```
          <location>
               <pp>765</pp>
               <column>
                    <colf>R</colf>
               </column>
          </location>
     </article>
</nplcit>
```

3) For NPL_TYPE = s
/nplcit/article/location/pp/ppf

```
<nplcit num="2" npl-type="s">
     <text>SCARDOVI, V., GENUS BIFIDOBACTERIUM., pages 1418 - 1434</text>
     <article>
          <author>
               <name>SCARDOVI, V.</name>
          </author>
          <serial>
               <sertitle>GENUS BIFIDOBACTERIUM.</sertitle>
          </serial>
          <location>
               <pp>
                    <ppf>1418</ppf>
                    <ppl>1434</ppl>
               </pp>
          </location>
     </article>
</nplcit>
```

4) For NPL_TYPE = d
/nplcit/online/location/pp

```
<nplcit num="1" npl-type="d">
     <text>SOVIET INVENTIONS ILLUSTRATED Week 8319, 22 June 1983 Derwent
Publications Ltd., London, GB; Page 9, AN 83-G3615K</text>
     <online>
          <edition>5</edition>
          <pubdate>19830622</pubdate>
          <vid>83</vid>
          <ino>19</ino>
          <absno>83-G3615K</absno>
          <location>
               <pp>9</pp>
          </location>
     </online>
     <source-doc>
          <document-id>
               <country>SU</country>
               <doc-number>939826</doc-number>
               <kind>A1</kind>
          </document-id>
     </source-doc>
</nplcit>
```

5) For NPL_TYPE = w
/nplcit/online/location/pp/ppf

```
<nplcit num="1" npl-type="w" extracted-xp="002552951">
```

```
     <text>A. G. STEWARD ET AL: "CATALYTIC CHAIN TRANSFER POLYMERISATION OF
FUNCTIONAL METHACRYLATES", INTERNET ARTICLE, 1998, pages 1 - 11, XP002552951,
Retrieved from the Internet
&lt;URL:http://www.warwick.ac.uk/fac/sci/Chemistry/polymers/downloads/stewardascm
1998.pdf&gt; [retrieved on 20091029]</text>
     <online>
          <author>
               <name>A. G. STEWARD ET AL</name>
          </author>
          <online-title>CATALYTIC CHAIN TRANSFER POLYMERISATION OF FUNCTIONAL
METHACRYLATES</online-title>
          <serial>
               <sertitle>INTERNET ARTICLE</sertitle>
          </serial>
          <pubdate>
               <sdate>1998</sdate>
          </pubdate>
          <location>
               <pp>
                    <ppf>1</ppf>
                    <ppl>11</ppl>
               </pp>
          </location>
     <avail>http://www.warwick.ac.uk/fac/sci/Chemistry/polymers/downloads/steward
ascm1998.pdf</avail>
          <srchdate>
               <date>20091029</date>
          </srchdate>
     </online>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.120 NPL_PAGE_LAST

Name: Last page of citation
Also Known As: n/a
Description: End of page range.
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Serial / Journal / Periodical citation (s)
- World Wide Web / Internet search citation (w)
Domain: up to 200 ASCII characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b
/nplcit/article/book/location/pp/ppl

```
<nplcit num="4" npl-type="b">
     <text>"ICI Handbook", pages: 1612 - 13</text>
     <article>
          <book>
               <book-title>ICI Handbook</book-title>
               <location>
                    <pp>
                         <ppf>1612</ppf>
                         <ppl>13</ppl>
                    </pp>
               </location>
          </book>
     </article>
</nplcit>
```

2) For NPL_TYPE = s
/nplcit/article/location/pp/ppl

```
<nplcit num="2" npl-type="s">
     <text>SCARDOVI, V., GENUS BIFIDOBACTERIUM., pages 1418 - 1434</text>
     <article>
          <author>
               <name>SCARDOVI, V.</name>
          </author>
          <serial>
               <sertitle>GENUS BIFIDOBACTERIUM.</sertitle>
          </serial>
          <location>
               <pp>
                    <ppf>1418</ppf>
                    <ppl>1434</ppl>
               </pp>
          </location>
     </article>
</nplcit>
```

3) For NPL_TYPE = w
/nplcit/online/location/pp/ppl

```
<nplcit num="1" npl-type="w" extracted-xp="002552951">
     <text>A. G. STEWARD ET AL: "CATALYTIC CHAIN TRANSFER POLYMERISATION OF
FUNCTIONAL METHACRYLATES", INTERNET ARTICLE, 1998, pages 1 - 11, XP002552951,
Retrieved from the Internet
&lt;URL:http://www.warwick.ac.uk/fac/sci/Chemistry/polymers/downloads/stewardascm
1998.pdf&gt; [retrieved on 20091029]</text>
     <online>
          <author>
               <name>A. G. STEWARD ET AL</name>
          </author>
          <online-title>CATALYTIC CHAIN TRANSFER POLYMERISATION OF FUNCTIONAL
METHACRYLATES</online-title>
          <serial>
               <sertitle>INTERNET ARTICLE</sertitle>
          </serial>
          <pubdate>
               <sdate>1998</sdate>
          </pubdate>
          <location>
               <pp>
                    <ppf>1</ppf>
                    <ppl>11</ppl>
               </pp>
          </location>
     <avail>http://www.warwick.ac.uk/fac/sci/Chemistry/polymers/downloads/steward
ascm1998.pdf</avail>
          <srchdate>
               <date>20091029</date>
          </srchdate>
     </online>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.121 NPL_PUBLISHER

Name: Publisher or name of host database
Also Known As: n/a
Description: Name of publisher (for book citations) or name of host / documentation database (for database citations).
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Derwent citation (d)
- Database citation (e)
- World Wide Web / Internet search citation (w)
Domain: up to 500 characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b
/nplcit/article/book/imprint/name

```
<nplcit num="3" npl-type="b">
     <text>NOBLE, W.C.: "The skin microflora and microbial skin disease", 2004,
CAMBRIDGE UNIV. PRESS</text>
     <article>
          <book>
               <author>
                    <name>NOBLE, W.C.</name>
               </author>
               <book-title>The skin microflora and microbial skin disease</book-
title>
               <imprint>
                    <name>CAMBRIDGE UNIV. PRESS</name>
                    <pubdate>2004</pubdate>
               </imprint>
          </book>
     </article>
</nplcit>
```

2) For NPL_TYPE = d, e, w
/nplcit/online/hosttitle

```
<nplcit num="1" npl-type="e" extracted-xp="002504853">
     <text>DATABASE CAPLUS [online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO,
US; MEIJERINK, J. ET AL: "Identification of olfactory stimulants for   Anopheles
gambiae from human sweat samples", XP002504853, retrieved from STN Database
accession no. 2000:446093</text>
     <online>
          <author>
               <name>MEIJERINK, J. ET AL</name>
          </author>
          <online-title>Identification of olfactory stimulants for   Anopheles
gambiae from human sweat samples</online-title>
          <hosttitle>CAPLUS</hosttitle>
          <imprint>
               <name>CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US</name>
          </imprint>
```

```
          <absno>2000:446093</absno>
          <avail>STN</avail>
     </online>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.122 NPL_PUBLN_DATE

Name: (Start) Date of the publication of the NPL
Also Known As: n/a
Description: Date (or year or month) of the publication of this Non-Patent Literature.
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Chemical Abstract citation (c)
- Biological abstract citation (i)
- Patent Abstracts of Japan (j)
- Serial / Journal / Periodical citation (s)
- Derwent citation (d)
- Database citation (e)
- World Wide Web / Internet search citation (w)

Domain: String with up to 8 digits:
Typical values are of the form yyyy, yyyy00, yyyymm or yyyyymmdd.
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b
/nplcit/article/book/imprint/pubdate

```
<nplcit num="3" npl-type="b">
     <text>NOBLE, W.C.: "The skin microflora and microbial skin disease", 2004,
CAMBRIDGE UNIV. PRESS</text>
     <article>
          <book>
               <author>
                    <name>NOBLE, W.C.</name>
               </author>
               <book-title>The skin microflora and microbial skin disease</book-
title>
               <imprint>
                    <name>CAMBRIDGE UNIV. PRESS</name>
                    <pubdate>2004</pubdate>
               </imprint>
          </book>
     </article>
</nplcit>
```

2) For NPL_TYPE = c, i, j
/nplcit/article/serial/pubdate

```
<nplcit num="1" npl-type="c" extracted-xp="002170177">
     <text>CHEMICAL ABSTRACTS, vol. 81, no. 18, 4 November 1974, Columbus, Ohio,
US; abstract no. 110011, BELOUSOV, V. YA.: &quot;New materials in gas and
petroleum industry machine building&quot; XP002170177</text>
     <article>
          <author>
               <name>BELOUSOV, V. YA.</name>
          </author>
          <atl>New materials in gas and petroleum industry machine building</atl>
          <serial>
               <sertitle/>
```

```
            <pubdate>19741104</pubdate>
            <vid>81</vid>
            <ino>18</ino>
        </serial>
        <absno>110011</absno>
    </article>
</nplcit>
```

3) For NPL_TYPE = s
/nplcit/article/serial/pubdate/sdate

```
<nplcit num="1" npl-type="s" extracted-xp="004519684">
    <text>LV J ET AL: "Controlled growth of three morphological structures of
magnesium hydroxide nanoparticles by wet precipitation method", JOURNAL OF
CRYSTAL GROWTH, ELSEVIER, AMSTERDAM, NL, vol. 267, no. 3-4, 1 July 2004 (2004-07-
01), pages 676 - 684, XP004519684, ISSN: 0022-0248</text>
    <article>
        <author>
            <name>LV J ET AL</name>
        </author>
        <atl>Controlled growth of three morphological structures of magnesium
hydroxide nanoparticles by wet precipitation method</atl>
        <serial>
            <sertitle>JOURNAL OF CRYSTAL GROWTH, ELSEVIER, AMSTERDAM,
NL</sertitle>
            <pubdate>
                <sdate>20040701</sdate>
            </pubdate>
            <vid>267</vid>
            <ino>3-4</ino>
            <issn>0022-0248</issn>
        </serial>
        <location>
            <pp>
                <ppf>676</ppf>
                <ppl>684</ppl>
            </pp>
        </location>
    </article>
</nplcit>
```

4) For NPL_TYPE = d
/nplcit/online/pubdate

```
<nplcit num="1" npl-type="d" extracted-xp="002139174">
    <text>SOVIET PATENTS ABSTRACTS Section PQ Week 8839, 9 November 1988 Derwent
Publications Ltd., London, GB; Class P56, AN 88276863, XP002139174</text>
    <online>
        <edition>6</edition>
        <pubdate>19881109</pubdate>
        <vid>88</vid>
        <ino>39</ino>
        <absno>88276863</absno>
        <location>
            <sersect>PQ</sersect>
        </location>
        <class>P56</class>
    </online>
    <source-doc>
```

```
          <document-id>
               <country>SU</country>
               <doc-number>1380915</doc-number>
               <kind>A1</kind>
          </document-id>
     </source-doc>
</nplcit>
```

5) For NPL_TYPE = e
/nplcit/online/imprint/pubdate

```
<nplcit num="5" npl-type="e" extracted-xp="002549869">
     <text>DATABASE CAPLUS [online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO,
US; 10 March 2009 (2009-03-10), WAN, JUNXI ET AL: "Ultrasonic preparation method
of magnesium hydroxide nanopowder", XP002549869, retrieved from STN Database
accession no. 2009:278478</text>
     <online>
          <author>
               <name>WAN, JUNXI ET AL</name>
          </author>
          <online-title>Ultrasonic preparation method of magnesium hydroxide
nanopowder</online-title>
          <hosttitle>CAPLUS</hosttitle>
          <imprint>
               <name>CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US</name>
               <pubdate>20090310</pubdate>
          </imprint>
          <absno>2009:278478</absno>
          <avail>STN</avail>
     </online>
</nplcit>
```

6) For NPL_TYPE = w
/nplcit/online/pubdate/sdate

```
<nplcit num="3" npl-type="w">
     <text>HUNTERLAB, APHA BACKGROUND, APPLICATIONS NOT, INSIGHT ON COLOR, vol.
8, no. 16, 16 November 1996 (1996-11-16), Retrieved from the Internet
&lt;URL:http://www.hunterlab.com/appnotes/anll 96br2.pdf.&gt;</text>
     <online>
          <serial>
               <sertitle>HUNTERLAB, APHA BACKGROUND, APPLICATIONS NOT, INSIGHT ON
COLOR</sertitle>
               <vid>8</vid>
               <ino>16</ino>
          </serial>
          <pubdate>
               <sdate>19961116</sdate>
          </pubdate>
          <avail>http://www.hunterlab.com/appnotes/anll 96br2.pdf.</avail>
     </online>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.123 NPL_PUBLN_END_DATE

Name:
Also Known As: n/a
Description: End date of the publication of this Non-Patent Literature.
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Serial / Journal / Periodical citation (s)
- World Wide Web / Internet search citation (w)
Domain: String with up to 8 digits:
Typical values are of the form yyyy, yyyy00, yyyymm or yyyyymmdd.
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = s
/nplcit/article/serial/pubdate/edate

```
<nplcit num="6" npl-type="s" extracted-xp="002503936">
     <text>SUGAI, MORIMITSU, IWSAKI, MORITA, WATANABE, KUBOTA: "Pungent qualities
of sanshool-related compounds evaluated by a sensory test and activation of rat
TRPV1", BIOSCIENCE, BIOTECNOLOGY AND BIOCHEMISTRY, vol. 69, no. 10, 2005 - 1951,
pages 1957, XP002503936</text>
     <article>
          <author>
               <name>SUGAI, MORIMITSU, IWSAKI, MORITA, WATANABE, KUBOTA</name>
          </author>
          <atl>Pungent qualities of sanshool-related compounds evaluated by a
sensory test and activation of rat TRPV1</atl>
          <serial>
               <sertitle>BIOSCIENCE, BIOTECNOLOGY AND BIOCHEMISTRY</sertitle>
               <pubdate>
                    <sdate>2005</sdate>
                    <edate>1951</edate>
               </pubdate>
               <vid>69</vid>
               <ino>10</ino>
          </serial>
          <location>
               <pp>
                    <ppf>1957</ppf>
               </pp>
          </location>
     </article>
</nplcit>
```

2) For NPL_TYPE = w
/nplcit/online/pubdate/edate

```
<nplcit num="2" npl-type="w" extracted-xp="002540162">
     <text>S. SENGUPTA, S. RAYANCHU, S. BANERJEE: "An Analysis of Wireless
Network Coding for Unicast Sessions: The Case for Coding-Aware Routing", INFOCOM
2007, 6 May 2007 (2007-05-06) - 12 May 2007 (2007-05-12), pages 1028 - 1036,
XP002540162, Retrieved from the Internet
&lt;URL:http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=04215706&gt;
[retrieved on 20090730]</text>
```

```
     <online>
          <author>
               <name>S. SENGUPTA, S. RAYANCHU, S. BANERJEE</name>
          </author>
          <online-title>An Analysis of Wireless Network Coding for Unicast
Sessions: The Case for Coding-Aware Routing</online-title>
          <serial>
               <sertitle>INFOCOM 2007</sertitle>
          </serial>
          <pubdate>
               <sdate>20070506</sdate>
               <edate>20070512</edate>
          </pubdate>
          <location>
               <pp>
                    <ppf>1028</ppf>
                    <ppl>1036</ppl>
               </pp>
          </location>
     <avail>http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=04215706</avail>
          <srchdate>
               <date>20090730</date>
          </srchdate>
     </online>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.124 NPL_PUBLN_ID

Name: Non-Patent Literature publication identification
Also Known As: n/a
Description: Surrogate key for Non-Patent Literature publications
Domain: 32 ASCII characters, or the digit 0
Default value: 0 (only for the dummy NPL record in table TLS214_NPL_PUBLN)

Source database: This value is the MD5 hash of the attribute NPL_BIBLIO. The only exception is a dummy record in table TLS214_NPL_PUBLN, which has an NPL_PUBLN_ID = 0.
Source field name: n/a

Comments
n/a

Modification history

Author of update - Date of update - Explanation of update
R. Heijna - 01-07-2005 - First version
R. Heijna - 21-11-2005 - Name and definition adapted
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 13-04-2012 - PATSTAT uses <refno> as surrogate key
M. Kracker - 01-04-2014 – Clarification in source field name and comment
M. Kracker - 01-10-2019 – NPL_PUBLN_ID can have more than 9 digits
M. Kracker - 01-04-2021 – The attribute has been completely re-defined: different domain, different values

6.125 NPL_TITLE1

Name: First title
Also Known As: n/a
Description: Title of article or article within the electronic resource.
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Chemical Abstract citation (c)
- Biological abstract citation (i)
- Serial / Journal / Periodical citation (s)
- Derwent citation (d)
- Database citation (e)
- World Wide Web / Internet search citation (w)

Domain: up to 1 000 characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b, c, i, s
/nplcit/article/atl

```
<nplcit num="1" npl-type="s" extracted-xp="009124942">
      <text>ASANO S ET AL: "Infection of herpes simplex virus type 2 mutant
lacking US3 induces apoptosis in the corneal epithelium of mice", IOVS, vol. 39,
no. 4, 15 March 1998 (1998-03-15), &amp;amp; ANNUAL MEETING OF THE ASSOCIATION FOR
RESEARCH IN VISION AND OPHTHALMOLOGY; FORT LAUDERDALE, FLORIDA, USA; MAY 10-15,
1998, pages S1064, XP009124942</text>
      <article>
            <author>
                  <name>ASANO S ET AL</name>
            </author>
            <atl>Infection of herpes simplex virus type 2 mutant lacking US3
induces apoptosis in the corneal epithelium of mice</atl>
            <serial>
                  <sertitle>IOVS</sertitle>
                  <imprint>
                        <address>
                              <text>&amp;amp; ANNUAL MEETING OF THE ASSOCIATION FOR
RESEARCH IN VISION AND OPHTHALMOLOGY; FORT LAUDERDALE, FLORIDA, USA; MAY 10-15,
1998</text>
                        </address>
                        <name/>
                  </imprint>
                  <pubdate>
                        <sdate>19980315</sdate>
                  </pubdate>
                  <vid>39</vid>
                  <ino>4</ino>
            </serial>
            <location>
                  <pp>
                        <ppf>S1064</ppf>
                  </pp>
            </location>
      </article>
</nplcit>
```

2) For NPL_TYPE = d, e, w
/nplcit/online/online-title

```
<nplcit num="6" npl-type="e" extracted-xp="002551016">
     <text>DATABASE CAPLUS [online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO,
US; 15 February 2007 (2007-02-15), ARITA, HIROAKI; FUKUDA, KAZUHIRO: "Gas-barrier
ceramic film laminates, resin substrates equipped with them, and organic
electroluminescent devices therewith", XP002551016, retrieved from STN Database
accession no. 2007:166400</text>
     <online>
          <author>
               <name>ARITA, HIROAKI; FUKUDA, KAZUHIRO</name>
          </author>
          <online-title>Gas-barrier ceramic film laminates, resin substrates
equipped with them, and organic electroluminescent devices therewith</online-
title>
          <hosttitle>CAPLUS</hosttitle>
          <imprint>
               <name>CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US</name>
               <pubdate>20070215</pubdate>
          </imprint>
          <absno>2007:166400</absno>
          <avail>STN</avail>
     </online>
</nplcit>
```

Comments
See also attribute NPL_TITLE2.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.126 NPL_TITLE2

Name: Second title
Also Known As: n/a
Description: Title of the book or serial.
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Chemical Abstract citation (c)
- Biological abstract citation (i)
- Patent Abstracts of Japan (j)
- Serial / Journal / Periodical citation (s)
- Derwent citation (d)
- World Wide Web / Internet search citation (w)

Domain: up to 1 000 characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b
/nplcit/article/book-title

```
<nplcit num="20" npl-type="b">
     <text>MURRAY: "Methods in Molecular Biology, Volume 7: Gene Transfer and
Expression Protocols", vol. 7, 1991, THE HUMANA PRESS, INC., article BAILEY ET
AL.: "Manipulation of Baculovirus Vectors", pages: 147 - 168</text>
     <article>
          <author>
               <name>BAILEY ET AL.</name>
          </author>
          <atl>Manipulation of Baculovirus Vectors</atl>
          <book>
               <author>
                    <name>MURRAY</name>
               </author>
               <book-title>Methods in Molecular Biology, Volume 7: Gene Transfer
and Expression Protocols</book-title>
               <imprint>
                    <name>THE HUMANA PRESS, INC.</name>
                    <pubdate>1991</pubdate>
               </imprint>
               <vid>7</vid>
               <location>
                    <pp>
                         <ppf>147</ppf>
                         <ppl>168</ppl>
                    </pp>
               </location>
          </book>
     </article>
</nplcit>
```

2) For NPL_TYPE = c, i, j, s
/nplcit/article/serial/sertitle

```
<nplcit num="1" npl-type="s">
```

```
     <text>HLB, J. SOC. COSMET. CHEM., vol. 1, pages 1949</text>
     <article>
          <serial>
               <sertitle>HLB, J. SOC. COSMET. CHEM.</sertitle>
               <vid>1</vid>
          </serial>
          <location>
               <pp>
                    <ppf>1949</ppf>
               </pp>
          </location>
     </article>
</nplcit>
```

3) For NPL_TYPE = d, w
/nplcit/online/serial/sertitle

```
<nplcit num="1" npl-type="w" extracted-xp="025543349">
     <text>SAITO H ET AL: "Cytotoxicity of chlorophenols to goldfish GFS cells
with the MTT and LDH assays", TOXICOLOGY IN VITRO, ELSEVIER SCIENCE, GB, vol. 8,
no. 5, 1 October 1994 (1994-10-01), pages 1107 - 1112, XP025543349, ISSN: 0887-
2333, [retrieved on 19941001]</text>
     <online>
          <author>
               <name>SAITO H ET AL</name>
          </author>
          <online-title>Cytotoxicity of chlorophenols to goldfish GFS cells with
the MTT and LDH assays</online-title>
          <serial>
               <sertitle>TOXICOLOGY IN VITRO, ELSEVIER SCIENCE, GB</sertitle>
               <vid>8</vid>
               <ino>5</ino>
               <issn>0887-2333</issn>
          </serial>
          <pubdate>
               <sdate>19941001</sdate>
          </pubdate>
          <location>
               <pp>
                    <ppf>1107</ppf>
                    <ppl>1112</ppl>
               </pp>
          </location>
          <srchdate>
               <date>19941001</date>
          </srchdate>
     </online>
</nplcit>
```

Comments
See also attribute NPL_TITLE1.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version
M. Kracker - 01-10-2018 – TITLE2 can occur also for NPL_TYPE = d

6.127 NPL_TYPE

Name: Type of the Non-Patent Literature
Also Known As: n/a
Description: Indicates the type of the Non-Patent Literature
Domain: 1 ASCII character:

For poor NPL citations (no rich NPL structure):
a Abstract citation of no specific kind

For articles:
b Book citation
c Chemical abstracts citation
i Biological abstract citation
j Patent Abstracts of Japan citation
s Serial / Journal / Periodical citation

For online citations:
d Derwent citation
e Database citation
w World Wide Web / Internet search citation

For the dummy entry:
space dummy

Default value: "a"
Note: all NPL citations with poor data content (in contrast to articles and online citations, which have a “rich" NPL structure) will have NPL_TYPE = "a”.
Source database: DOCDB
Source field name:

```
<publn_nplcit>
      <nplcit num="1" npl-type="s" extracted-xp="055067747">
            <text>JINEK M. ET AL: "A PROGRAMMABLE DUAL-RNA-GUIDED DNA
ENDONUCLEASE IN ADAPTIVE BACTERIAL IMMUNITY (SUPPLEMENTARY MATERIAL)", SCIENCE,
vol. 337, no. 6096, 17 August 2012 (2012-08-17), XP055067747</text>
            <article>
                  <author>
                        <name>JINEK M. ET AL</name>
                  </author>
                  <atl>A PROGRAMMABLE DUAL-RNA-GUIDED DNA ENDONUCLEASE IN
ADAPTIVE BACTERIAL IMMUNITY (SUPPLEMENTARY MATERIAL)</atl>
                  <serial>
                        <sertitle>SCIENCE</sertitle>
                        <pubdate>
                              <sdate>20120817</sdate>
                        </pubdate>
                        <vid>337</vid>
                        <ino>6096</ino>
                  </serial>
            </article>
      </nplcit>
</publn_nplcit>
```

Source sub-field identifier: n/a

Comments: In case of data inconsistencies (diverging NPL types for the same NPL) in the source database, priority is given to the type which occurs most frequently in rich citations.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2016 - First version

6.128 NPL_VOLUME

Name: Volume ID / number
Also Known As: n/a
Description: ID or number of the volume.
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Book citation (b)
- Chemical Abstract citation (c)
- Biological abstract citation (i)
- Patent Abstracts of Japan (j)
- Serial / Journal / Periodical citation (s)
- Derwent citation (d)
- World Wide Web / Internet search citation (w)

Domain: up to 50 characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = b
/nplcit/article/book/vid

```
<nplcit num="3" npl-type="b">
     <text>"Bergey's manual of systematic Bacteriology.", vol. 2, 1986, WILLIAMS
&amp;amp; WILKINS, pages: 635</text>
     <article>
          <book>
               <book-title>Bergey's manual of systematic Bacteriology.</book-title>
               <imprint>
                    <name>WILLIAMS &amp;amp; WILKINS</name>
                    <pubdate>1986</pubdate>
               </imprint>
               <vid>2</vid>
               <location>
                    <pp>
                         <ppf>635</ppf>
                    </pp>
               </location>
          </book>
     </article>
</nplcit>
```

2) For NPL_TYPE = c, i, j, s
/nplcit/article/serial/vid

```
<nplcit num="1" npl-type="s">
     <text>HLB, J. SOC. COSMET. CHEM., vol. 1, pages 1949</text>
     <article>
          <serial>
               <sertitle>HLB, J. SOC. COSMET. CHEM.</sertitle>
               <vid>1</vid>
          </serial>
               <location>
               <pp>
                    <ppf>1949</ppf>
               </pp>
```

```
          </location>
     </article>
</nplcit>
```

3) For NPL_TYPE = d
/nplcit/online/vid

```
<nplcit num="1" npl-type="d" extracted-xp="002510294">
     <text>DATABASE WPI Week 198718, Derwent Publications Ltd., London, GB; AN
1987-125518, XP002510294</text>
     <online>
          <edition>0</edition>
          <vid>1987</vid>
          <ino>18</ino>
          <absno>1987-125518</absno>
     </online>
     <source-doc>
          <document-id>
               <country>JP</country>
               <doc-number>S6267014</doc-number>
               <kind>A</kind>
          </document-id>
     </source-doc>
</nplcit>
```

4) For NPL_TYPE = w
/nplcit/online/serial/vid

```
<nplcit num="1" npl-type="w" extracted-xp="025543349">
     <text>SAITO H ET AL: "Cytotoxicity of chlorophenols to goldfish GFS cells
with the MTT and LDH assays", TOXICOLOGY IN VITRO, ELSEVIER SCIENCE, GB, vol. 8,
no. 5, 1 October 1994 (1994-10-01), pages 1107 - 1112, XP025543349, ISSN: 0887-
2333, [retrieved on 19941001]</text>
     <online>
          <author>
               <name>SAITO H ET AL</name>
          </author>
          <online-title>Cytotoxicity of chlorophenols to goldfish GFS cells with
the MTT and LDH assays</online-title>
          <serial>
               <sertitle>TOXICOLOGY IN VITRO, ELSEVIER SCIENCE, GB</sertitle>
               <vid>8</vid>
               <ino>5</ino>
               <issn>0887-2333</issn>
          </serial>
          <pubdate>
               <sdate>19941001</sdate>
          </pubdate>
          <location>
               <pp>
                    <ppf>1107</ppf>
                    <ppl>1112</ppl>
               </pp>
          </location>
          <srchdate>
               <date>19941001</date>
          </srchdate>
     </online>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.129 NUTS

Name: NUTS region, region
Also Known As: Nomenclature of Territorial Units for Statistics
Description: This attribute contains the NUTS code as defined by Eurostat
Domain: 2 - 5 ASCII characters
In table TLS206_PERSON: NUTS will be empty for countries not covered by NUTS codes
Default value: empty (only in table TLS206_PERSON)
Source database: made available by ECOOM (K.U. LEUVEN)
Source field name: n/a
Source sub-field identifier: n/a

Comments:
The regionalisation procedure is performed on all person addresses of EP patents available in the previous PATSTAT edition, and adopts the new NUTS 2013 codes for 37 countries (28 EU members, 5 member candidates (Albania, Macedonia, Montenegro, Serbia and Turkey) and 4 EFTA countries (Iceland, Liechtenstein, Norway and Switzerland

See table TLS904_NUTS for the label of the NUTS codes.

The processing of the NUTS regions starts after PATSTAT data is released. Therefore the result will be available sometime later. Consequently, additions and changes introduced by the current PATSTAT edition are not included in the current edition but will be included in the next edition of PATSTAT.

The methodology for regionalisation is described in
http://ec.europa.eu/eurostat/documents/3859598/5916785/KS-RA-11-008-EN.PDF.

Since the PATSTAT 2020 Autumn Edition the NUTS codes of some person records (in TLS206_PERSON) are enhanced by values from the REGPAT January 2020 database provided by the OECD. For more details see http://oe.cd/ipstats .
These enhanced records are identified by a NUTS_LEVEL with the value 4. OECD is also using NUTS version 2013 (plus minor additions), except for the United Kingdom, where the NUTS version 2010 is used for the London area.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2016 - First version
M. Kracker - 01-04-2018 – Comments changed.
M. Kracker - 01-10-2020 – Comments changed (Enhancement with OECD REGPAT data)

6.130 NUTS_LABEL

Name: Name of the NUTS region code
Also Known As: n/a
Description: The name of the region according to NUTS (Nomenclature of Territorial Units for Statistics), version 2013 in original language
Domain: up to 250 characters
Default value: n/a

Source database: made available by ECOOM (K.U. LEUVEN), which is based on public data from Eurostat
Source field name: n/a
Source sub-field identifier: n/a

Comments:
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2016 - First version

6.131 NUTS_LEVEL

Name: Level of NUTS region code
Also Known As: n/a
Description: Indicates the level of the regionalisation code in attribute NUTS (according to NUTS - Nomenclature of Territorial Units for Statistics)
Domain: 1-digit number with values 0, 1, 2, 3, 4 or 9.
Value 0 indicates that the NUTS code identifies a state.
Values 1, 2 and 3 are the official NUTS levels.
Value 4 is equivalent to NUTS level 3 but indicates that the NUTS code has been provided by OECD's REGPAT database.
Value 9 indicates that no NUTS code has been assigned.

Default value: 9

Source database: made available by ECOOM (K.U. LEUVEN)
Source field name: n/a
Source sub-field identifier: n/a

Comments:
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2016 - First version
M. Kracker - 01-10-2020 – Value 4 has been added to the domain

6.132 OECD_MEMBER

Name: Member of the Organisation for Economic Co-operation and Development
Also Known As: n/a
Description: Indicates whether this country/territory is a member state of the OECD
Domain: 1 ASCII character: Y or space

Y	If a country/territory is member of the OECD.
space	otherwise

Default value: n/a
Source database: http://www.oecd.org/about/membersandpartners/
Source field name: n/a
Source sub-field identifier: n/a
Comments: This field indicates the OECD members at the time of the production of the PATSTAT edition. Depending on the time range you need to analyse, you may want to exclude “newer” OECD members (e.g. Colombia).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version

6.133 ONLINE_AVAILABILITY

Name: Online availability
Also Known As: n/a
Description: Access information for the online citation. This may contain information about the web URL, the online database, the FTP address, email etc. **. .**
This attribute may only be populated for these NPL types (see attribute NPL_TYPE):
- Database citation (e)
- World Wide Web / Internet search citation (w)
Domain: up to 500 characters
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = e, w
/nplcit/online/avail

```
<nplcit num="5" npl-type="e" extracted-xp="002549869">
     <text>DATABASE CAPLUS [online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO,
US; 10 March 2009 (2009-03-10), WAN, JUNXI ET AL: "Ultrasonic preparation method
of magnesium hydroxide nanopowder", XP002549869, retrieved from STN Database
accession no. 2009:278478</text>
     <online>
          <author>
               <name>WAN, JUNXI ET AL</name>
          </author>
          <online-title>Ultrasonic preparation method of magnesium hydroxide
nanopowder</online-title>
          <hosttitle>CAPLUS</hosttitle>
          <imprint>
               <name>CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US</name>
               <pubdate>20090310</pubdate>
          </imprint>
          <absno>2009:278478</absno>
          <avail>STN</avail>
     </online>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version
M. Kracker - 01-10-2017 – Domain extended to 500 characters

6.134 ONLINE_CLASSIFICATION

Name: Online classification
Also Known As: n/a
Description: One or more Derwent classes**,** as used in the Derwent citations.
This attribute may only be populated for this NPL type (see attribute NPL_TYPE):
- Derwent citation (d)

Domain: up to 30 ASCII characters;
Each Derwent class consists of 3 characters: 1 letter followed by 2 digits. Multiple Derwent classes are separated by a comma.
Examples: "D22" or "E32,M25"
More information about the Derwent classes (= DWPI classification system) can be found in https://clarivate.com/derwent/wp-content/uploads/sites/3/dlm_uploads/2019/08/DWPI-Classification-Guide-2020.pdf .

Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = d
/nplcit/online/class

```
<nplcit num="1" npl-type="d">
     <text>DATABASE WPI Week 200952, 7 July 2009 Derwent Publications Ltd.,
London, GB; Class E32 M25, AN 2009-L51362</text>
     <online>
          <edition>1</edition>
          <pubdate>20090707</pubdate>
          <vid>2009</vid>
          <ino>52</ino>
          <absno>2009-L51362</absno>
          <class>E32 M25</class>
     </online>
     <source-doc>
          <document-id doc-id="276140550">
               <country>BR</country>
               <doc-number>PI0705592</doc-number>
               <kind>A2</kind>
          </document-id>
     </source-doc>
</nplcit>
```

Note: It assumed that all Derwent classes are contained in a single <class> element. Spaces in DOCDB's XML will to be replaced by commas in PATSTAT.

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

M. Kracker - 01-10-2017 - ONLINE_CLASSIFICATION may hold more than 1 Derwent class

6.135 ONLINE_SEARCH_DATE

Name: Online search date
Also Known As: n/a
Description: Date of search or retrieval.
This attribute may only be populated for this NPL type (see attribute NPL_TYPE):
- World Wide Web / Internet search citation (w)
Domain: String with up to 8 digits: Typical values are of the form yyyy, yyyymm or yyyyymmdd.
Default value: empty

Source database: DOCDB
Source field name:

1) For NPL_TYPE = w
/nplcit/online/srchdate/date

```
<nplcit num="4" npl-type="w" extracted-xp="002503935">
     <text>SUCCAR, MITCHELL, VAUGHAN: "Actions of N-arachidonyl-glycine in a rat
inflammatory pain model", 30 August 2007 (2007-08-30), XP002503935, Retrieved
from the Internet &lt;URL:http://www.molecularpain.com/content/3/1/24&gt;
[retrieved on 20081114]</text>
     <online>
          <author>
               <name>SUCCAR, MITCHELL, VAUGHAN</name>
          </author>
          <online-title>Actions of N-arachidonyl-glycine in a rat inflammatory
pain model</online-title>
          <pubdate>
               <sdate>20070830</sdate>
          </pubdate>
          <avail>http://www.molecularpain.com/content/3/1/24</avail>
          <srchdate>
               <date>20081114</date>
          </srchdate>
     </online>
</nplcit>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.136 ORGANISATION_FLAG

Name: Organisation flag
Also Known As: n/a
Description: Indicates that a two-letter code indicates an organisation (and not a state / territory)
Domain: 1 ASCII character: Y or space

Y	if this two-letter code indicates an organisation (and not a state / territory)
space	otherwise

Default value: n/a
Source database: WIPO standard ST.3
Source field name: n/a
Source sub-field identifier: n/a

Comments: This attribute is useful if data has to be displayed on a map

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version
M. Kracker - 01-04-2020 – Replaced attribute STATE_INDICATOR

6.137 PARENT_APPLN_ID

Name: Application identification of the earlier application
Also Known As: n/a
Description: Surrogate key of the application which was the basis for the continuation application
Domain: Number 1 … 999 999 999
Only earlier applications for which a continuation is filed with the same authority (domestic). E.g. the country in the priority-claim is the same as the country in the application-reference. Clearly self-claimers are to be ignored. The **continuation** is published with an INID-code in the 60-series (WIPO ST.9) (plus inner priority, INID (23) as used by DE). The relevant case is case # 6 from section 4.6 "Relation Types".
The PARENT_APPLN_ID is taken from the APPLN_ID allocated in PATSTAT for the earlier application. All of the applications must have been collected from the DOCDB before this logic can be used.

Default value: n/a
Source database: DOCDB, PATSTAT

Source field name

```
<priority-claim sequence="2" data-format="docdb" status="A">
      <country>US</country>
      <doc-number>9885602</doc-number>
      <kind>A</kind>
      <date>20020314</date>
      <priority-linkage-type>3</priority-linkage-type>
      <priority-active-indicator>N</priority-active-indicator>
</priority-claim>
<priority-claim sequence="2" data-format="epodoc">
      <doc-number>US20020098856</doc-number>
</priority-claim>
<priority-claim sequence="1" data-format="original">
      <doc-number>9885602/doc-number>
</priority-claim>
```

With `<country>US</country> <doc-number>9885602</doc-number> <kind>A</kind>` in DOCDB the corresponding application in PATSTAT is determined (via APPLN_AUTH, APPLN_NR and APPLN_KIND) and the value of APPLN_ID for this corresponding application is the PARENT_APPLN_ID

If there is no corresponding application in PATSTAT it should be created, see section 4.4 "Application replenishment".

Source sub-field identifier
n/a
Comments
n/a
Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 03-05-2005 - First version
R. Heijna - 20-07-2005 - Source field definition improved
R. Heijna - 07-07-2005 - Value zero for the physical model
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML

6.138 PARTY_NEW

Name: New party
Also Known As: n/a
Description: The name of the new party. It may also contain address data, concatenated and separated by "," or ";"
Domain: Up to 1 000 characters
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-details/parties/parties-details/party/name

```
<legal-event providing-office="EP" date-added="20111020" date-previous-
exchange="20111020" sequence-number="4">
     <event-date>20111019</event-date>
     <event-code>RIN1</event-code>
     <event-details>
          <event-description event-description-type="original">ERFINDER
(KORR.)</event-description>
          <event-description lang="en">INVENTOR (CORRECTION)</event-description>
          <parties party-type="inventor" sequence-number="1">
               <parties-details>
                    <party>
                         <name>KANG, CHUL-KYU</name>
                    </party>
               </parties-details>
          </parties>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.139 PARTY_OLD

Name: Old party
Also Known As: n/a
Description: Unstructured text containing "former owners" as well as any other text; can contain multiple names and addresses
Domain: Up to 1 000 characters
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-details/parties/text

```
<legal-event event-type="REG" providing-office="HU" date-added="20140821" date-
previous-exchange="20140823" sequence-number="46">
     <event-date>20140728</event-date>
     <event-code>GB9C</event-code>
     <event-details>
          <event-description event-description-type="original">JOGUTODLAS</event-
description>
          <event-description lang="en">SUCCESSION IN TITLE</event-description>
          <parties party-type="owner" sequence-number="1">
               <parties-details>
                    <party>
                         <name>DELPHI INTERNATIONAL OPERATIONS LUXEMBOURG S.A,
LU</name>
                    </party>
               </parties-details>
               <text>FORMER OWNER(S): DELPHI TECHNOLOGIES HOLDING S.A.R.L.,
LU</text>
          </parties>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.140 PARTY_SEQ_NR

Name: Party sequence number
Also Known As: n/a
Description: Sequence number of the party
Domain: number 0 .. 50 (but currently 0 or 1; see comment below)
Default value: 0

Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-details/parties/@sequence-number

```
<legal-event providing-office="EP" date-added="20111020" date-previous-
exchange="20111020" sequence-number="4">
     <event-date>20111019</event-date>
     <event-code>RIN1</event-code>
     <event-details>
          <event-description event-description-type="original">ERFINDER
(KORR.)</event-description>
          <event-description lang="en">INVENTOR (CORRECTION)</event-description>
          <parties party-type="inventor" sequence-number="1">
               <parties-details>
                    <party>
                         <name>KANG, CHUL-KYU</name>
                    </party>
               </parties-details>
          </parties>
     </event-details>
</legal-event>
```

Comments
Currently (2021 Autumn Edition) it is assumed that there will only be at most 1 party for an event, so PARTY_SEQ_NR will always be 0 or 1.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.141 PARTY_TYPE

Name: Party type
Also Known As: n/a
Description: Type of the party: owner, inventor, representative, opponent or licensee.
Domain: 3 ASCII characters or empty
OWN owner
INV inventor
REP representative
OPP opponent
LIC licensee
OTH other
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-details/parties/@party-type

```
<legal-event providing-office="EP" date-added="20111020" date-previous-
exchange="20111020" sequence-number="4">
     <event-date>20111019</event-date>
     <event-code>RIN1</event-code>
     <event-details>
          <event-description event-description-type="original">ERFINDER
(KORR.)</event-description>
          <event-description lang="en">INVENTOR (CORRECTION)</event-description>
          <parties party-type="inventor" sequence-number="1">
               <parties-details>
                    <party>
                         <name>KANG, CHUL-KYU</name>
                    </party>
               </parties-details>
          </parties>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.142 PAT_CITN_SEQ_NR

Name: Sequence number of the patent citation
Also Known As: n/a
Description: Number for a patent citation in the series of patent citations for one publication/"origin of citation" combination
Domain: Number 0 … about 1000
Default value: 0
Source database: Computed from PATSTAT. It is a sequential number for each patent citation, regardless whether the patent citation is referring to a patent publication (CITED_PAT_PUBLN_ID > 0) or patent application (CITED_APPLN_ID > 0). The numbering starts with 1 for each origin of citations (CITN_ORIGIN).
The PAT_CITN_SEQ_NR will be set to 0 when the citation is not a patent citation, but an NPL (non-patent literature) citation.
Source field name: n/a

Source sub-field identifier: n/a
Comments
The PAT_CITN_SEQ_NR attribute does ***not*** indicate the order of appearance of patent citations.

The sequence numbers start at 1 for each origin of the citations.

The sequence number identifies all patent citations, i.e. it does not distinguish between citations of publications (CITED_PAT_PUBLN_ID) and citations of applications (CITED_APPLN_ID).

See also attributes NPL_CITN_SEQ_NR and CITN_ID.

Modification history

Author of update - Date of update - Explanation of update
R. Heijna - 04-05-2005 - First version
R. Heijna - 01-07-2005 - Specifically for patent citations
R. Heijna - 15-07-2005 - For PL as well as NPL citations
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 08-05-2013 - Clarification in comment
M. Kracker - 01-12-2015 – Changed processing instructions and comments

6.143 PAT_PUBLN_ID

Name: Patent publication identification
Also Known As: n/a
Description: Surrogate key for patent publications
Domain: Number 0 … 999 999 999
Range 1 - from 1 to 900 000 000 is used for publications of standard applications (= non-artificial applications)
Range 2 - from 900 000 001 to 999 999 999 for artificial publication references created in PATSTAT for those cited publications which do not themselves have a publication reference in DOCDB (see section 4.5 “Publication replenishment”).
Default value: 0

Source database: DOCDB (for Range 1), PATSTAT (for Range 2)

Source field name
For Range 1:
<exchange-document country="EP" doc-number="0681755" kind="B1" **doc-id="300943156"** date-publ="19960904" family-id="21747543" is-representative="NO" date-of-last-exchange="20150206" date-added-docdb="19960831" originating-office="EP">

For Range 2: A unique number is generated for each unique combinations of the alternate key (PUBLN_AUTH, PUBLN_NR, PUBLN_KIND, PUBLN_DATE).
Source sub-field identifier: n/a

Comments
Note: For reasons of database consistency, there must be a dummy publication with a PAT_PUBLN_ID value of 0.

Within range 1 (1 to 900 000 000) this key will remain stable, i.e. it will not change between PATSTAT editions. For details see section 4.3.2 “Stable IDs”.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 04-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 0104-2014 – Changed comment; PATPUBLN_ID need not be sequential
M. Kracker - 0104-2015 – IDs of non-artificial publications are now taken from DOCDB

6.144 PERSON_ADDRESS

Name: Person Address
Also Known As: Correspondence address
Description: All address elements of the person apart from the country. Example: street, city, postal code.
Domain: Up to 1.000 characters
Default value: empty
Source database: see attribute PERSON_NAME

For DOCDB data:
Source field name and Source sub-field identifier:
see attribute ADDRESS_FREEFORM

For EP Register data:
Source field name
The attributes ADDRESS_1, ADDRESS_2, ADDRESS_3, ADDRESS_4 and ADDRESS_5 of table TLS226_PERSON_ORIG are concatenated with a comma.

For USPTO data of published applications and published grants:
Source field name:
The attributes STREET, CITY, ZIP_CODE and STATE of table TLS226_PERSON_ORIG are concatenated.
Alternatively – if available- the attributes ADDRESS_1, ADDRESS_2 and ADDRESS_3 are concatenated.

Regardless of the source, each source data element of this attribute is cleaned:

- Leading and trailing spaces are removed
- whitespace characters (tabs, line feed, carriage return, ...) are replaced by a space
- multiple spaces are reduced to a single space

Comments
Address data in DOCDB is only available for a few authorities with scattered coverage: EP, IT, WO, CA, FI, AT and for older GB and IE documents. Therefore better-quality address for EP and US patents is taken from other sources: The EPO address data is sourced from the EP Register. The USPTO address data is sourced from the USPTO publication files on USPTO's website.

In PATSTAT Online due to data privacy reasons, the PERSON_ADDRESS has been emptied for all persons who might be a natural person (e. g. all inventors, or where the SECTOR attribute contains “INDIVIDUAL” or “UNKNOWN” or is empty.)

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 01-12-2004 - First version
R. Heijna - 31-05-2005 - Applicants and Inventors integrated
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-10-2013 - For EP changed source to EP Register; Changed description of source
M. Kracker - 05-10-2014 – Comment updated

M. Kracker - 01-05-2015 – Comment updated
M. Kracker - 01-10-2015 – Section “Source field name” for EP Register data is amended
M. Kracker - 01-04-2016 – USPTO source may use ADDRESS_1 – ADDRESS_3

6.145 PERSON_CTRY_CODE

Name: Person country code
Also Known As: cc, country, territory
Description: Country part of the correspondence address of the person or business
Domain: 2 characters (A-Z), according to WIPO ST.3 (plus minor additions)or spaces
Exception: in case of bad data it may be any characters (e. g. "UK" is not ST.3 compliant, but should be "GB")
Default value: spaces
Source database: DOCDB, EP Register, and USPTO data of published applications and published grants
Source field name: see attribute PERSON_NAME

DOCDB data:
Source field name

```
<inventors>
      <inventor sequence="1" data-format="docdb">
            <inventor-name>
                  <name>STACY N SMITH</name>
            </inventor-name>
            <residence>
                  <country>US</country>
            </residence>
      </inventor>
      <inventor sequence="1" data-format="docdba">
            <inventor-name>
                  <name>STACY N. SMITH</name>
            </inventor-name>
            <address>
                  <text>305 Cottonwood Lane, NC 27540 Holly Springs,
UNITED STATES OF AMERICA (USA)</text>
            </address>
      </inventor>
      <inventor sequence="1" data-format="original">
            <inventor-name>
                  <name>Stacy N. Smith</name>
            </inventor-name>
      </inventor>
</inventors>


<applicants>
      <applicant sequence="1" data-format="docdb">
            <applicant-name>
                  <name>ERICSSON INC</name>
            </applicant-name>
            <residence>
                  <country>US</country>
            </residence>
      </applicant>
      <applicant sequence="1" data-format="docdba">
            <applicant-name>
                  <name>ERICSSON INC.</name>
            </applicant-name>
            <address>
                  <text>7001 Development Drive, 27709-3969 Research
Triangle Park,UNITED STATES OF AMERICA (USA)</text>
            </address>
```

```
      </applicant>
      <applicant sequence="1" data-format="original">
            <applicant-name>
                  <name>Ericsson Inc.</name>
            </applicant-name>
      </applicant>
</applicants>
```

Source sub-field identifier

```
data-format="docdb"
```

For EP Register data:

Source field name

```
<parties>
      <applicants change-gazette-num="2000/29">
            <applicant app-type="applicant" designation="all" sequence="1">
                  <addressbook>
                        <name>Seidel, Helmut</name>
                        <address>
                              <address-1>Fliederstrasse 19</address-1>
                              <address-2>65396 Walluf</address-2>
                              <country>DE</country>
                        </address>
                  </addressbook>
                  <nationality>
                        <country/>
                  </nationality>
                  <residence>
                        <country/>
                  </residence>
            </applicant>
      </applicants>
      <inventors change-gazette-num="2000/29">
            <inventor sequence="01">
                  <addressbook>
                        <name>Franta, Georg</name>
                        <address>
                              <address-1>Ulrich-Rapp-Strasse 18</address-1>
                              <address-2>87634 Obergünzburg</address-2>
                              <country>DE</country>
                        </address>
                  </addressbook>
            </inventor>
            <inventor sequence="02">
                  <addressbook>
                        <name>Dojan, Viktor</name>
                        <address>
                              <address-1>Ludwig-Strecker-Strasse 5</address-1>
                              <address-2>55129 Mainz</address-2>
                              <country>DE</country>
                        </address>
                  </addressbook>
            </inventor>
      </inventors>
```

Source sub-field identifier: n/a

For USPTO data of published applications and published grants:

Source field name:

```
<applicant sequence="00" app-type="applicant-inventor" designation="us-only">
```

```
      <addressbook>
            <last-name>Eckhoff</last-name>
            <first-name>Philip A.</first-name>
            <address>
                  <city>Bellevue</city>
                  <state>WA</state>
                  <country>US</country>
            </address>
      </addressbook>
      <nationality>
            <country>omitted</country>
      </nationality>
      <residence>
            <country>US</country>
      </residence>
</applicant>
```

Comments
The country code as well as the full name of the country are usually indicated as "the country". Note: Only for about 50% of the inventors their country code is known.

Note that the EPO does not receive the Country Code value with the Japanese data which is loaded into DOCDB; for this reason there are no PERSON_CTRY_CODEs in PATSTAT for Japanese documents.

This code is copied from the 'standard' DOCDB table and added to the 'bypass' data, matching on the application id of authority, number and kind code and inventor sequence number or applicant sequence number.

Country code does not necessarily indicate the "Nationality" of inventor or applicant.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 20-12-2004 - First version
R. Heijna - 31-05-2005 - Applicants and Inventors integrated
D. Lingua - 13-02-2008 - Comment and other fields modified
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 15-05-2013 - Added exception to Domain; For EP applications changed source to EP Register; Changed description of source
M. Kracker - 01-04-2019 – Domain extended

6.146 PERSON_ID

Name: Person identification
Also Known As: n/a
Description: Surrogate key based on the elements in the alternate primary key of table TLS206_PERSON
Domain: Number 1 … 999 999 999
Default value: n/a
Source database: PATSTAT
Source field name
Computed based on PERSON_NAME, PERSON_NAME_ORIG_LG, PERSON_ADDRESS and PERSON_CTRY_CODE in PATSTAT. Allocate a surrogate key PERSON_ID for each combination of these fields. Upper case and lower case are considered equal. E.g. “James Bond” is considered to be the same person name as “JAMES BOND”.

Source sub-field identifier: n/a
Comments
Sequential number unique for each unique combination of the elements in the candidate primary key.

Persons are the legal or physical persons that have a relation with the patent granting procedure. Currently included are applicants and inventors.

This key will normally remain stable, i.e. it will not change between PATSTAT editions. However, in exceptional cases some values of PERSON_ID might change. For details see section 4.3.2 “Stable IDs”.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 15-04-2005 - First version
R. Heijna - 31-05-2005 - Applicant integrated with Inventor
M. Kracker - 01-10-2013 - Clarification in description
M. Kracker - 15-10-2015 - Clarification in comment
M. Kracker - 01-04-2015 – PERSON_NAME_ORIG_LG added as source field

6.147 PERSON_NAME

Name: Person name
Also Known As: n/a
Description: Name of the Applicant or Inventor
Domain: Up to 500 characters
Default value: empty
Source database:
1) EP Register for EP patent applications

2) OECD patents database for US data post 1976-01-01 up to and including November 15th, 2005 for Published Grants.

3) PATSTAT weekly file extracts from USPTO website for Published Grants from November 22nd, 2005 until today, and for Published Applications from September 29th, 2005 until today.

4) Inventor & Applicant names for USPTO Published Applications from March 1st, 2001 to September 22nd, 2005 from DOCDB, data-format="docdba".

5) all other names from DOCDB, data-format="docdba".

For DOCDB data:
Source field name and Source sub-field identifier: see attribute NAME_FREEFORM

For EP Register data:
Source field name and Source sub-field identifier: see attribute NAME_FREEFORM

For USPTO data of published applications and published grants:
Source field name:
The attributes LAST_NAME, FIRST_NAME and MIDDLE_NAME of table TLS226_PERSON_ORIG are concatenated. The results, depending on the availability of the data, are like

- Kennedy, John F
- Kennedy, John
- Kennedy

Regardless of the source, each source data element of this attribute is cleaned:

- Leading and trailing spaces are removed
- whitespace characters (tabs, line feed, carriage return, ...) are replaced by a space
- multiple spaces are reduced to a single space

Comments
See also Business Rules in section 5.6 ”TLS206_PERSON: Person”.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 01-12-2004 - First version
R. Heijna - 15-04-2005 - Size, source, comments updated

R. Heijna - 31-05-2005 - Applicants and Inventors integrated
J. Rollinson - 18-04-2006 - Comments extended, source field and domain updated
D. Lingua - 15-10-2008 - Comments extended
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-10-2013 - Source fields for EP Register added; Comments updated
M. Kracker - 15-10-2014 - Comments updated
M. Kracker - 01-04-2016 – Domain extended to 500 characters

6.148 PERSON_NAME_ORIG_LG

Name: Person name in original language
Also Known As: non-transliterated person name
Description: Non-transliterated person name in original character set, which may be Japanese, Chinese, Korean, Arabic, Cyrillic, etc. characters
Domain: up to 500 characters
Default value: empty
Source database: DOCDB
Source field name

Example from publication JP 2015011369 A, published 2015-01-19

```
<exch:parties>
      <exch:applicants>
            <exch:applicant sequence="1" data-format="docdb">
            <exch:applicant-name>
                  <name>NEC COMMUNICATION SYST</name>
            </exch:applicant-name>
            </exch:applicant>
            <exch:applicant sequence="1" data-format="docdba">
                  <exch:applicant-name>
                        <name>NEC COMMUN SYST LTD</name>
                  </exch:applicant-name>
            </exch:applicant>
            <exch:applicant sequence="1" data-format="original">
                  <exch:applicant-name>
                        <name>日本電気通信システム株式会社</name>
                  </exch:applicant-name>
            </exch:applicant>
      </exch:applicants>
      <exch:inventors>
            <exch:inventor sequence="1" data-format="docdb">
                  <exch:inventor-name>
                        <name>KUDO KENTARO</name>
                  </exch:inventor-name>
            </exch:inventor>
            <exch:inventor sequence="1" data-format="docdba">
                  <exch:inventor-name>
                        <name>KUDO KENTARO</name>
                  </exch:inventor-name>
            </exch:inventor>
            <exch:inventor sequence="1" data-format="original">
                  <exch:inventor-name>
                        <name>工藤　健太郎</name>
                  </exch:inventor-name>
            </exch:inventor>
      </exch:inventors>
</exch:parties>
```

Comments
If there is no original name for a person available, then

- in table TLS226_PERSON_ORIG this attribute remains empty
- in table TLS206_PERSON this attribute is replenished with the value of PERSON_NAME

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2019 - First version

6.149 PERSON_ORIG_ID

Name: Key for the unmodified person data record
Also Known As: n/a
Description: Number which uniquely identifies a row in the TLS226_PERSON_ORIG table
Domain: Number 1 … 999 999 999
Default value: n/a
Source database: PATSTAT
Source field name: n/a
Source sub-field identifier: n/a
Comments: This key will remain stable, i.e. it will not change between PATSTAT editions. However, in exceptional cases some values of PERSON_ORIG_ID might change. For details see section 4.3.2 “Stable IDs”.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version

6.150 PRIOR_APPLN_ID

Name: Application identification of claimed Paris Convention priority
Also Known As: n/a
Description: Surrogate key of an application of which the priority is claimed under the Paris convention
Domain: Number, 1 ... 999 999 999
Only "pure" priorities i.e. those according to the Paris convention and published with an INID-code in the 30-series (WIPO ST.9). The relevant case is case # 2 from section 4.6 "Relation Types".
Default value: n/a
Source database: DOCDB, PATSTAT
Source field name

```
<priority-claims>
      <priority-claim sequence="1" data-format="docdb">
            <document-id>
                  <country>DE</country>
                  <doc-number>10331291</doc-number>
                  <kind>A</kind>
                  <date>20030710</date>
            </document-id>
      <priority-active-indicator>Y</priority-active-indicator>
</priority-claim>
```

With

```
<country>DE</country>
<doc-number>10331291</doc-number>
<kind>A</kind>
```

in DOCDB the corresponding priority application in PATSTAT is determined (via APPLN_AUTH, APPLN_NR and APPLN_KIND) and the value of APPLN_ID of this priority application will be assigned to PRIOR_APPLN_ID.

If there is no corresponding application in PATSTAT it should be created, see section 4.4 "Application replenishment".
Source sub-field identifier
n/a
Comments
n/a
Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 22-04-2005 - First version
R. Heijna - 20-07-2005 - Source field definition improved
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML

6.151 PRIOR_APPLN_SEQ_NR

Name: Sequence number of claimed priority
Also Known As: n/a
Description: Number indicating the place in the list of priorities claimed in the application.
Domain: Number, 1... about 500
Default value: n/a
Source database: DOCDB
Source field name:

```
<...>
<priority-claims>
      <priority-claim sequence="1" data-format="docdb">
            <document-id>
                  <country>DE</country>
                  <doc-number>10331291</doc-number>
                  <kind>A</kind>
                  <date>20030710</date>
            </document-id>
      <priority-active-indicator>Y</priority-active-indicator>
</priority-claim>
```

If an application is claiming itself as a priority, then this priority is not stored in PATSTAT. So if a priority-claim element is the same as the application-reference, the application is claiming itself as a priority. These are normally the last priority in the priority-claims list of DOCDB.

This means that the sequence numbers of any subsequent priorities claimed by this application must be reduced by 1. See the rules for PRIOR_APPLN_ID to see which priorities are to be ignored.

Source sub-field identifier
n/a
Comments
The sequence number is assigned based on the sequence in which the priorities have been provided by the supplier.
For US data - where sequence is extremely important with continuations/divisions/continuations in part - the sequence numbers is on filing-date descending. Earliest filing date last.
Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 22-12-2004 - First version
D. Lingua - 05-06-2009 - Added comments
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML

6.152 PSN_ID

Name: ID for the PATSTAT Standardised Name
Also Known As: n/a
Description: PSN_NAMEs which have been harmonised according to the Univ. Leuven harmonisation procedure have a unique PSN_ID for each unique PSN_NAME. Multiple rows may have the same PSN_ID, if multiple person names in the person table have been harmonised into a single PSN name.
PSN_NAMEs which have *not* been harmonised this way have a unique PSN_ID for each (un-harmonised) PERSON_NAME.
Domain: Number 1 … 999 999 999
Default value: n/a
Source database: Computed from data made available by ECOOM (K.U. LEUVEN)
Not all PSN_NAMEs have undergone the harmonisation process (cf. attribute PSN_LEVEL).

- For PSN_NAMEs which have been created during the harmonisation process the unique PSN_ID for each PSN_NAME is in the range 1 ... 100 000 000
- For PSN_NAMEs which have *not* been created during the harmonisation process but which just have been replenished by the PERSON_NAME the number is computed as “PERSON_ID + 100 000 000”.

Source field name: n/a
Source sub-field identifier: n/a
Comments:
The processing of the PATSTAT standardised name starts as soon as PATSTAT data is released. Typically the result will be available 3-4 months afterwards. Consequently, additions and changes introduced by the current PATSTAT edition are not harmonised in the current edition but will be harmonised in the next edition of PATSTAT.

See also comment of attribute PSN_NAME.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 15-10-2014 – Comment updated
M. Kracker - 01-04-2015 – Comment updated
M. Kracker - 01-04-2016 – Attribute renamed (was: HRM_L2_ID); Comment updated
M. Kracker - 01-04-2017 – Clarifications in Description and Source Database

6.153 PSN_LEVEL

Name: Harmonisation level of PATSTAT standardised name
Also Known As: n/a
Description: This attribute indicates for each name in PSN_NAME the level of harmonisation which has been applied
Domain: Number 0 … 2

0: No harmonisation has taken place
(PSN_NAME is the same as attribute PERSON_NAME)
1: Automated harmonisation only has been applied
2: Automated harmonisation plus manual refinement have been applied.

Default value: n/a
Source database: made available by ECOOM (K.U. LEUVEN)
Source field name: n/a
Source sub-field identifier: n/a
Comments:
The processing of the PATSTAT standardised name starts as soon as PATSTAT data is released. Typically the result will be available 3-4 months afterwards. Consequently, additions and changes introduced by the current PATSTAT edition are not harmonised in the current edition but will be harmonised in the next edition of PATSTAT.
See also comment of attribute PSN_NAME.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 15-10-2014 – Comment updated
M. Kracker - 01-04-2016 – Attribute renamed (was: HRM_LEVEL); Comment updated

6.154 PSN_NAME

Name: PATSTAT standardised name
Also Known As: n/a
Description: PATSTAT standardised name.
The attribute is populated for all persons. Names of persons which have not been harmonised are just copied from the attribute PERSON_NAME.
Domain: Up to 500 characters
Default value: n/a
Source database: made available by ECOOM (K.U. LEUVEN)
Source field name: n/a
Source sub-field identifier: n/a
Comments:
The processing of the PATSTAT standardised name starts as soon as PATSTAT data is released. Typically the result will be available 3-4 months afterwards. Consequently, additions and changes introduced by the current PATSTAT edition are not harmonised in the current edition but will be harmonised in the next edition of PATSTAT.
The PATSTAT standardised names are the results of an approach to standardise the original name (https://www.ecoom.be/en/data-collections/patstat-enhancements). It is done in an automated way with additional manual refinements.

Background papers on the production of the PSN_NAME:

i. Du Plessis, M., Van Looy, B., Song, X & Magerman, T. (2009) Data Production Methods for Harmonized Patent Indicators: Assignee sector allocation. EUROSTAT Working Paper and Studies, Luxembourg.
ii. Magerman T, Grouwels J., Song X. & Van Looy B. (2009). Data Production Methods for Harmonized Patent Indicators: Patentee Name Harmonization. EUROSTAT Working Paper and Studies, Luxembourg.
iii. Peeters B., Song X., Callaert J., Grouwels J., Van Looy B. (2009). Harmonizing harmonized patentee names: an exploratory assessment of top patentees. EUROSTAT Working paper and Studies, Luxembourg.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 15-10-2014 – Updated description and comment
M. Kracker - 01-04-2016 – Attribute renamed (was: HRM_L2);
Domain extended to 500 characters; Comment updated

6.155 PSN_SECTOR

Name: Sector of the applicant
Also Known As: n/a
Description: This is a by-product of the PATSTAT standardised name harmonisation effort: Applicants may have been assigned to one or more sectors, like company, government or non-profit organization, university or hospital. If the sector of an applicant cannot be determined, then the sector is UNKNOWN. If a person (e.g. a person who is only an inventor, but not an applicant) is not assigned a sector, then this field is empty.
So this column may contain zero, one or more of these keywords:

INDIVIDUAL
COMPANY
UNKNOWN
GOV
NON-PROFIT
UNIVERSITY
HOSPITAL

This list of keywords may change.
Domain: Up to 50 ASCII characters
Default value: empty
Source database: made available by ECOOM (K.U. LEUVEN)
Source field name: n/a
Source sub-field identifier: n/a
Comments:
The processing of the PATSTAT standardised name starts as soon as PATSTAT data is released. Typically the result will be available 3-4 months afterwards. Consequently, additions and changes introduced by the current PATSTAT edition are not harmonised in the current edition but will be harmonised in the next edition of PATSTAT.

See also comment of attribute PSN_NAME.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 15-10-2014 – Comment updated
M. Kracker - 01-04-2016 – Attribute renamed (was: SECTOR); Comment updated

6.156 PUBLN_AUTH

Name: Publication Authority
Also Known As: Publishing office
Description: Patent Authority that issued the publication of the application
Domain: 2 characters (A-Z) according to WIPO ST.3 or spaces
Exception: in case of bad data it may be any characters
Default value: spaces
Source database: DOCDB
Source field name:

1) Standard publication reference (PAT_PUBLN_ID between 0 and 900 000 000):

```
<publication-reference data-format="docdb">
      <document-id lang="en">
            <country>US</country>
            <doc-number>2007011914</doc-number>
            <kind>A1</kind>
            <date>20070118</date>
      </document-id>
</publication-reference>
```

2) Artificial publications from cited references (PAT_PUBLN_ID between 900 00 001 and 999 999 999):

```
<references-cited>
      <citation srep-phase="SEA" sequence="1">
            <patcit>
                  <document-id>
                        <country>US</country>
                        <doc-number>3380531</doc-number>
                        <kind>A</kind>
                  </document-id>
            </patcit>
            <category>A</category>
      </citation>
```

Source sub-field identifier
data-format="docdb"
Comments
Take all the publication-references in DOCDB into PATSTAT.
For all artificial publications which were themselves artificially created for those cited publications, where the cited publications are not registered in DOCDB as publications: use the authority (country) of the cited publication: `<country>US</country>`

Check if the cited publication has a publication-reference in DOCDB and if not, then create an artificial publication and an artificial application.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 04-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 15-05-2013 - Added exception to Domain

6.157 PUBLN_CLAIMS

Name: Indicator of the number of claims in the given publication
Also Known As: n/a
Description: This indicator provides the number of claims that has been attributed to the specific publication. Currently only available for certain EP and US publications.
Domain: Number 0 … about 1.000;
0 means

- that a publication contains no claims (e. g. EP publications of kind A3, A8, B8 and non-republished EuroPCT applications)
  or
- that the number of claims is not known (e. g. US-publications published on or before 1974; publications not from EP or US)

Default value: 0
Source database: Special delivery files for EP and US data only
Source field name:

There are two separate sources for US and EP data. Specific formats are used in each case.

1) US data: relates to granted patents only (A documents until 2000, B1 or B2 documents afterwards) which were published on or after 1975-01-01

The backfile published by the USPTO has this format:

- Columns 1-7: US Patent document
  I.e. issued patent to which the other information in the record applies (character field)
- Columns 9-12: Number of Claims
  If information for this field is missing the field is marked with a period (".") (integer field)
- Columns 14-17: Number of Drawing Figures
  If information for this field is missing, the field is marked with a period (".") (integer field)
- Columns 19-22: Number of Submitted Drawing Sheets
  If information for this field is missing, the field is marked with a period (".") (integer field)

Usage example:

```
7585234   24    4    3
7585235    8   18    7
7585236   42   23   13
```

Only the information in columns 1-7 and 9-12 is used, the remaining is ignored.

The information in column 1-7 is mapped to publication data in PATSTAT, where
PUBLN_AUTH is US
PUBLN_NR is the content of column 1-7
PUBLN_KIND is "A" until 2000, "B1" or "B2" from 2001

The content of column 9-12 is used to populate element PUBLN_CLAIMS.

2) EP data: relates to both published applications (kind code "A") from 1978 and granted patents (kind code "B") from 1980.

Data was extracted from the EPO publications XML, in this format:

```
EP publication number;kind code;publication date;number of claims
1123811;A2;20010816;17
1124248;A2;20010816;20
1123812;A2;20010816;34
```

The information needs to be mapped to publication data in PATSTAT as follows:
PUBLN_AUTH is EP.
PUBLN_NR is the content of "EP publication number".
PUBLN_KIND is the content of "kind code".

The content of "number of claims" is used to populate element PUBLN_CLAIMS.

Source sub-field identifier: n/a
Comments
In a minority of cases for EP B (European granted patents) publications, multiple sets of claims are published, each set applying to a specific group of designated states. For the sake of simplification, only the highest number of claims has been considered.

Warning: The number of claims will be "0" for all EP A documents originating from a PCT published in English, French or German (so called "Euro-PCTs"). For all these Euro-PCT documents, as the EPO does not republish the application (by recognising the PCT publication as being sufficient), the claim count for the EP document will be equal to "0" as there is no real EP A publication. For those Euro-PCT documents whose original PCT language is not English, French or German, there is a new publication in one EPO official language and thus the claim count is available.

Modification history
Author of update - Date of update - Explanation of update
D. Lingua - 04-08-2011 - First version
D. Lingua - 13-04-2012 - Update on EP B documents (1980 to 2005) and warning
D. Lingua - 25-09-2012 - Update on US claim coverage
M. Kracker - 01-10-2013 - Clarification for value 0
M. Kracker - 15-10-2015 - Clarification for coverage (US starting from 1975-01-01; EP)

6.158 PUBLN_DATE

Name: Publication date
Also Known As: n/a
Description: Date on which the publication was made available to the public
Domain: Date
Default value: 9999-12-31
Source database: DOCDB
Source field name

```
<publication-reference data-format="docdb">
      <document-id lang="en">
            <country>US</country>
            <doc-number>2007011914</doc-number>
            <kind>A1</kind>
            <date>20070118</date>
      </document-id>
</publication-reference>
```

With `country`, `doc-number` and `kind` in `document-id` in `patcit` in `citation` in `references-cited` in DOCDB the corresponding publication in PAT_PUBLN in PATSTAT is determined (via PUBLN_AUTH, PUBLN_NR and PUBLN_KIND). The value of PUBLN_DATE for this publication is the value of date in `document-id`. If it is an invalid date or empty, then use 9999-12-31.

Source sub-field identifier
`data-format="docdb"`
Comments
n/a
Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 04-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML

6.159 PUBLN_FIRST_GRANT

Name: Identifier of the granting publication of an application
Also Known As: n/a
Description: Indication that the publication can be considered as the first publication of grant of a given application
Domain: 1 ASCII character: Y or N:
N - this publication is **not** the first publication of a grant.
Y - this publication is the first publication of a grant.
Default value: n/a
Source database: DOCDB
Source field name: in tag `<date-of-public-availability>` when categories `<printed-with-grant> or <not-printed-with-grant>` are given

Case 1) **`<printed-with-grant>`**

```
<exchange-document country="ES" doc-number="2340887" kind="T3" date-
publ="20100610" family-id="38220640" is-representative="YES" date-of-last-
exchange="20100610" date-added-docdb="20100601" originating-office="EP"
status="C">
            <bibliographic-data>
                  <publication-reference data-format="docdb">
                        <document-id lang="es">
                              <country>ES</country>
                              <doc-number>2340887</doc-number>
                              <kind>T3</kind>
                              <date>20100610</date>
                        </document-id>
                  </publication-reference>
                  <publication-reference data-format="epodoc">
                        <document-id lang="es">
                              <doc-number>ES2340887T</doc-number>
                        </document-id>
                  </publication-reference>
...
                  <dates-of-public-availability>
                        <printed-with-grant>
                              <document-id>
                                    <date>20100610</date>
                              </document-id>
                        </printed-with-grant>
                  </dates-of-public-availability>
            </bibliographic-data>
```

Case 2) **`<not-printed-with-grant>`**

```
<exchange-document country="HK" doc-number="21891" kind="A" date-publ="19910328"
family-id="26321336" is-representative="YES" date-of-last-exchange="20100610"
date-of-previous-exchange="20001030" date-added-docdb="20000401" originating-
office="EP" status="A">
            <bibliographic-data>
                  <publication-reference data-format="docdb">
                        <document-id>
                              <country>HK</country>
                              <doc-number>21891</doc-number>
                              <kind>A</kind>
                              <date>19910328</date>
```

```
                        </document-id>
                    </publication-reference>
                    <publication-reference data-format="epodoc">
                        <document-id>
                            <doc-number>HK21891</doc-number>
                        </document-id>
                    </publication-reference>
...
                    <dates-of-public-availability>
                        <not-printed-with-grant>
                            <document-id>
                                <date>19910328</date>
                            </document-id>
                        </not-printed-with-grant>
                    </dates-of-public-availability>
                </bibliographic-data>
```

Source sub-field identifier: n/a

Comments
This indicator provides a somewhat simplistic view to identify the first publication of grant. It is based on the DOCDB XML element `<date-of-public-availability>` and will have a value "N" if this tag does not contain any of the two categories listed above. It will have the value "Y" if the tag contains one of the two categories listed above.
A value "N" is also given in case the element `<date-of-public-availability>` is not present.
In case there are multiple publications of a grant, the earlier publication only is given the "Y" indicator (first publication of grant).

Annex IV to the DOCDB ST36 Layout Description lists a concordance of this element with the WIPO ST.30 (http://www.wipo.int/export/sites/www/standards/en/pdf/03-30-01.pdf); codes "450" and "470" are the key to identify the granting publication.

The publication date of the granting publication can be considered as being the date of grant. However, exception exists, like the Austrian utility models which are granted 2 months before being published as a U1 publication.

Note: Some offices (MX, BR, AR, JP, ..) do not (always) publish granted patents but just issue a legal event. So looking at the legal event codes in PATSTAT Legal (table TLS231_INPADOC_LEGAL_EVENT) can reveal additional grants. The attribute GRANTED of table TLS201_APPLN takes both granting publications and legal event into account to deduce whether an application has been granted or not (cf. attribute GRANTED).

Although the EPO has taken great care in analysing the grant information, this process is the result of interpretations and assumptions for which no responsibility whatsoever can be accepted.

Modification history
Author of update Date of update - Explanation of update
D. Lingua - 23-02-2009 - First version
D. Lingua - 14-06-2010 - Changed source to DOCDB XML
M. Kracker - 01-04-2014 – Comment extended

M. Kracker - 01-10-2017 – Comment extended
M. Kracker - 01-10-2018 – Domain changed from 0/1 to N/Y; Comment changed.

6.160 PUBLN_KIND

Name: Kind of Publication
Also Known As: n/a
Description: Publication kind attributed by the Patent Authority issuing the publication
Domain: Up to 2 ASCII characters, as laid down in the "Kind Code concordance list" for databases within the EPO in column "DOCDB" available on https://www.epo.org/searching-for-patents/data/coverage/regular.html . See also "Comments" below.

Default value: n/a
Source database: DOCDB
Source field name
1) Standard publication reference (PAT_PUBLN_ID between 0 and 900 000 000):

```
<publication-reference data-format="docdb">
      <document-id lang="en">
            <country>US</country>
            <doc-number>2007011914</doc-number>
            <kind>A1</kind>
            <date>20070118</date>
      </document-id>
</publication-reference>
```

2) Artificial publications from cited references (PAT_PUBLN_ID between 900 00 001 and 999 999 999):

```
<references-cited>
      <citation srep-phase="SEA" sequence="1">
            <patcit>
                  <document-id>
                        <country>US</country>
                        <doc-number>3380531</doc-number>
                        <kind>A</kind>
                  </document-id>
            </patcit>
            <category>A</category>
      </citation>
```

Source sub-field identifier
`data-format="docdb"`

Comments
For all artificial publications created for those cited publications, where the cited publications are not registered in DOCDB as publications: use the authority (country) of the cited publication and the publication kind code as cited. Because of this, a substantial number of PUBLN_KIND codes will not occur in the DOCDB "Kind Code concordance list". Example: the data base contains more than 2 000 US publications with PUBLN_KIND code B. We assume this might be B1, B2, B3, ... but we have kept the kind code "B" as originally cited.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 04-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
D. Lingua - 11-10-2011 - Updated figures to October 2011 edition

6.161 PUBLN_LG

Name: Publication language
Also Known As: n/a
Description: Language of the publication
Domain: 2 ASCII characters, according to ISO language codes (ISO 639-1) or spaces
Default value: spaces
Source database: DOCDB
Source field name:

```
<publication-reference data-format="docdb">
      <document-id lang="en">
            <country>US</country>
            <doc-number>2007011914</doc-number>
            <kind>A1</kind>
            <date>20070118</date>
      </document-id>
</publication-reference>
```

If country = ‘DE’, then PUBLN_LG = 'DE'.

Source sub-field identifier
data-format="docdb"
Comments
Present in about 10% of cases only (NB not always necessary, e.g. DE publications are always in German)
Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 04-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML

6.162 PUBLN_NR

Name: Publication number
Also Known As: n/a
Description: Number given by the Patent Authority issuing the publication
Domain: Up to 15 ASCII characters (since April 2013 without leading spaces)
Default value: n/a
Source database: DOCDB
Source field name
1) standard publication reference:

```
<publication-reference data-format="docdb">
      <document-id lang="en">
            <country>US</country>
            <doc-number>2007011914</doc-number>
            <kind>A1</kind>
            <date>20070118</date>
      </document-id>
</publication-reference>
```

2) Publications from cited references:

```
<references-cited>
      <citation srep-phase="SEA" sequence="1">
            <patcit>
                  <document-id>
                        <country>US</country>
                        <doc-number>3380531</doc-number>
                        <kind>A</kind>
                  </document-id>
            </patcit>
            <category>A</category>
      </citation>
```

Source sub-field identifier
n/a
Comments
Most but not all offices give the same publication number to all publications of a given application. Exceptions are for example JP, CN and KR.

Note that the publication “number” is not necessarily numeric but may contain letters. Leading zeros might be relevant.

Modification history
Author of update - Date of update - Explanation of update
R. Heijna - 04-05-2005 - First version
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML
M. Kracker - 01-04-2019 - Comment amended

6.163 PUBLN_NR_ORIGINAL

Name: Publication number in original format
Also Known As: Original publication number
Description: Publication number in original format as provided by the supplier. It is assumed that the number is as printed on the publication.
The availability and the format of the original publication number depend on the publishing authority.
Domain: Up to 100 characters
Default value: empty
Source database: DOCDB
Source field name

1) Source for the standard (= non-artificial) publications:

```
<exch:publication-reference data-format="original">
      <document-id>
            <doc-number>07691110</doc-number>
      </document-id>
</exch:publication-reference>
```

If DOCDB does not provide an original publication number, then PUBLN_NR_ORIGINAL will contain an empty string.

2) For all artificial publications the attribute PUBLN_NR_ORIGINAL will contain an empty string.

Source sub-field identifier
data-format="original"
Source codes
n/a
Comments
This attribute is useful to combine publication data of PATSTAT with another publication data set which also contains the original publication number.
On average, only about 20% of all publications do have an original publication number.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2016 - First version

6.164 RECEIVING_OFFICE

Name: Receiving office
Also Known As: n/a
Description: Office where the international application was filed. Empty in case of regional or national filings.
Domain: Up to 2 ASCII characters (A-Z), according to WIPO ST.3
Default value: Spaces in case of non-PCT applications (attribute APPLN_KIND <> “W”)
Source database: DOCDB
Source field name

```
<application-reference data-format="docdb" is-representative="N">
      <document-id>
                  <country>DE</country>
                  <doc-number>10331291</doc-number>
                  <kind>W</kind>
                  <date>20030710</date>
```

Source sub-field identifier
data-format="docdb"

Comments
APPLN_KIND = “W” indicates an international application (= PCT application).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2018 - First version

6.165 REF_DOC_AUTH

Name: The publication authority of the referenced document.
Also Known As: n/a
Description: The publication authority of the referenced document. It is not indicated whether the referenced document is an application or a publication.
Domain: 2 ASCII characters (A-Z), according to WIPO ST.3
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/event-reference/event-ref-document/country

```
<legal-event event-type="REG" providing-office="DE" date-added="20120102" date-
previous-exchange="20120105" sequence-number="14">
     <event-date>20111229</event-date>
     <event-date-effective>20111229</event-date-effective>
     <event-code>R096</event-code>
     <event-details>
          <event-description event-description-type="original">VEROEFFENTLICHUNG
EINES HINWEISES AUF DIE EP-PATENTERTEILUNG DURCH DAS DPMA</event-description>
          <event-description lang="en">DPMA PUBLICATION OF MENTIONED EP PATENT
GRANT</event-description>
          <event-reference>
               <event-ref-document>
                    <country>DE</country>
                    <doc-number>602010000345</doc-number>
               </event-ref-document>
          </event-reference>
     </event-details>
</legal-event>
```

Comments
See Business Rules in the table description of TLS231_INPADOC_LEGAL_EVENT (section 0).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.166 REF_DOC_DATE

Name: Date of the referenced document.
Also Known As: n/a
Description: The date (application or publication date) of the referenced document. It is not indicated whether the referenced document is an application or a publication.
Domain: Date
Default value: 9999-12-31

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
1) /legal-status-document/legal-event/event-details/event-reference/event-ref-document/date
or
2) /legal-status-document/legal-event/event-details/event-reference/event-ref-kind/date

1)

```
<legal-event providing-office="EP" date-added="20110428" date-previous-
exchange="20110428" sequence-number="13">
     <event-date>20110428</event-date>
     <event-code>REF</event-code>
     <event-details>
          <event-description event-description-type="original">ENTSPRICHT</event-
description>
          <event-description lang="en">CORRESPONDS TO:</event-description>
          <event-reference>
               <event-ref-document>
                    <country>DE</country>
                    <doc-number>602010000011</doc-number>
                    <kind>P</kind>
                    <date>20110428</date>
               </event-ref-document>
          </event-reference>
     </event-details>
</legal-event>
```

2)
no example given

Comments
See Business Rules in the table description of TLS231_INPADOC_LEGAL_EVENT (section 5.24).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.167 REF_DOC_KIND

Name: Kind code of document
Also Known As: n/a
Description: The kind code of the referenced document. It is *not* indicated whether the referenced document is an application or a publication.
Domain: Up to 2 ASCII
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
1) /legal-status-document/legal-event/event-details/event-reference/event-ref-document/kind
or
2) /legal-status-document/legal-event/event-details/event-reference/event-ref-kind/kind

1)

```
<legal-event event-type="REG" providing-office="AT" date-added="20131218" date-
previous-exchange="20131221" sequence-number="18">
     <event-date>20131215</event-date>
     <event-date-effective>20131215</event-date-effective>
     <event-code>REF</event-code>
     <event-details>
          <event-description event-description-type="original">NENNUNG DER E-
NUMMER (EP PATENT TRITT IN AT NATIONALE PHASE EIN)</event-description>
          <event-description lang="en">REFERENCE TO AT NUMBER (EP PATENT ENTERS
AUSTRIAN NATIONAL PHASE)</event-description>
          <event-reference>
               <event-ref-document>
                    <country>AT</country>
                    <doc-number>642754</doc-number>
                    <kind>T</kind>
               </event-ref-document>
          </event-reference>
          </event-details>
</legal-event>
```

2)

```
<legal-event providing-office="EP" date-added="20110929" date-previous-
exchange="20110929" sequence-number="1">
     <event-date>20110928</event-date>
     <event-code>AK</event-code>
     <event-details>
          <event-description event-description-type="original">BENANNTE
VERTRAGSSTAATEN</event-description>
          <event-description lang="en">DESIGNATED CONTRACTING STATES:</event-
description>
          <event-reference>
               <event-ref-kind>
                    <kind>A1</kind>
               </event-ref-kind>
          </event-reference>
          <designated-states>
               <country>AT</country>
               <country>BE</country>
..
               <country>SM</country>
               <country>TR</country>
          </designated-states>
```

```
    </event-details>
</legal-event>
```

Comments

See Business Rules in the table description of TLS231_INPADOC_LEGAL_EVENT (section 0).

Modification history

Author of update - Date of update - Explanation of update

M. Kracker - 01-04-2017 - First version

6.168 REF_DOC_NR

Name: Serial number of the referenced document.
Also Known As: n/a
Description: Serial number of the referenced document. It is not indicated whether the referenced document is an application or a publication.
Domain: up to 20 ASCII characters. May contain letters and leading zeros.
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/event-reference/event-ref-document/doc-number

```
<legal-event event-type="REG" providing-office="DE" date-added="20120102" date-
previous-exchange="20120105" sequence-number="14">
     <event-date>20111229</event-date>
     <event-date-effective>20111229</event-date-effective>
     <event-code>R096</event-code>
     <event-details>
          <event-description event-description-type="original">VEROEFFENTLICHUNG
EINES HINWEISES AUF DIE EP-PATENTERTEILUNG DURCH DAS DPMA</event-description>
          <event-description lang="en">DPMA PUBLICATION OF MENTIONED EP PATENT
GRANT</event-description>
          <event-reference>
               <event-ref-document>
                    <country>DE</country>
                    <doc-number>602010000345</doc-number>
               </event-ref-document>
          </event-reference>
     </event-details>
</legal-event>
```

Comments
See Business Rules in the table description of TLS231_INPADOC_LEGAL_EVENT (section 0).

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.169 REF_DOC_TEXT

Name: Free text reference to a document.
Also Known As: n/a
Description: Free text reference to a document. It is not indicated whether the referenced document is an application or a publication.
Domain: up to 1 000 characters
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/event-reference/text

```
<legal-event providing-office="EP" date-added="20030130" date-previous-
exchange="20030101" sequence-number="11">
    <event-date>19950131</event-date>
    <event-code>EAL</event-code>
    <event-details>
        <event-description event-description-type="original">SE: EUROPEISKT
PATENT GAELLANDE I SVERIGE</event-description>
        <event-description lang="en">SE: EUROPEAN PATENT IN FORCE IN
SWEDEN</event-description>
        <event-reference>
            <text>86100760.7</text>
        </event-reference>
    </event-details>
</legal-event>
```

Comments
See Business Rules in the table description of TLS231_INPADOC_LEGAL_EVENT (section 5.24 ”TLS231_INPADOC_LEGAL_EVENT: Legal event”).
This field is only populated if there is not enough information to properly populate the fields REF_DOC_AUTH, REF_DOC_NR, REF_DOC_KIND or REF_DOC_DATE.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.170 REG_PHASE

Name: Indicator whether the application *is* or *has been* in the regional phase
Also Known As: n/a
Description: Indicates that an application *is* or *has been* in the regional phase.
Domain: 1 ASCII character

Y	Yes
N	No
space	not known (In case of uncertain interpretations; used very little or not at all)

Default value: N

Source database: PATSTAT
Source field name: Derived from tables TLS201_APPLN, TLS211_PAT_PUBLN and the "Kind code concordance list" (Excel format, see https://www.epo.org/searching-for-patents/helpful-resources/data/tables/regular.html)

Y if the APPLN_KIND <> W
and (APPLN_AUTH is a regional office
or (APPLN_AUTH is a member of a regional office
and the PUBLN_KIND code indicates that the patent publication is the result of a regional phase)
);
N otherwise

Source sub-field identifier: n/a

Comments
For explanation and disclaimer see attribute INT_PHASE in section 6.80.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2016 - First version

6.171 REINSTATE_COUNTRY

Name: Reinstatement country or territory
Also Known As: office of reinstatement
Description: Office of the application which has been reinstated. Contains the application authority and "WO" for PCT applications.
Domain: 2 ASCII characters (A-Z), according to WIPO ST.3
Default value: n/a

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/notification-of-reinstatement/@country

```
<legal-event providing-office="EP" date-added="20150505" date-previous-
exchange="20150509" sequence-number="38">
     <event-date>20150430</event-date>
     <event-code>PGRI</event-code>
     <event-details>
          <event-description event-description-type="original">POSTGRANT: PATENT
REINSTATED IN CONTRACTING STATE</event-description>
          <event-description lang="en">POSTGRANT: PATENT REINSTATED IN CONTRACTING
STATE</event-description>
          <notification-of-reinstatement country="SE">
               <date-patent-reinstated>20150303</date-patent-reinstated>
          </notification-of-reinstatement>
     </event-details>
</legal-event>
```

Comments
This attribute is populated if and only if EVENT_CODE = "PGRI".

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.172 REINSTATE_DATE

Name: Date of reinstatement
Also Known As: n/a
Description: Date when the reinstatement of a patent became effective
Domain: Date
Default value: 9999-12-31

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/notification-of- reinstatement/date-patent-reinstated

```
<legal-event providing-office="EP" date-added="20150505" date-previous-
exchange="20150509" sequence-number="38">
     <event-date>20150430</event-date>
     <event-code>PGRI</event-code>
     <event-details>
          <event-description event-description-type="original">POSTGRANT: PATENT
REINSTATED IN CONTRACTING STATE</event-description>
          <event-description lang="en">POSTGRANT: PATENT REINSTATED IN CONTRACTING
STATE</event-description>
          <notification-of-reinstatement country="SE">
               <date-patent-reinstated>20150303</date-patent-reinstated>
          </notification-of-reinstatement>
     </event-details>
</legal-event>
```

Comments
This attribute is populated if and only if EVENT_CODE = "PGRI ".

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.173 REINSTATE_TEXT

Name: Additional information about a reinstatement
Also Known As: n/a
Description: Additional information in free form text about the reinstatement of an application.
Domain: up to 1 000 characters
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/notification-of-reinstatement/text

Comments
This attribute is only populated if EVENT_CODE = "PGRI".
Note: Currently (2021 Autumn Edition) this field is never populated.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.174 RELEVANT_CLAIM

Name: Relevant claim
Also Known As: n/a
Description: Claim of the examined application to which a citation with a (set of) citation categorie(s) refers
Domain: Number between 0 and 1 000
Default value: 0, indicating that the relevant claim is not known
Source database: DOCDB
Source field name

```
<exch:citation cited-phase="SEA" cited-date="20120112" srep-office="EP" sequence="1">
	<patcit num="1" dnum="WO2010141409A1" dnum-type="publication number">
		<document-id doc-id="329547194">
			<country>WO</country>
			<doc-number>2010141409</doc-number>
			<kind>A1</kind>
			<name>CRONIN MICHAEL D [US], et al</name>
			<date>20101209</date>
		</document-id>
	</patcit>
	<rel-passage>
		<passage>
			<para>108</para>
			<figure>23A,23B</figure>
		</passage>
		<category>XP</category>
		<rel-claims>1-4,6,8,9,11</rel-claims>
		<category>I</category>
		<rel-claims>5</rel-claims>
	</rel-passage>
	<category>XPI</category>
</exch:citation>
```

Comments
RELEVANT_CLAIM is only populated with a non-default value if DOCDB contains the “rich” structure of a citation. For “poor” citations RELEVANT_CLAIM will always be 0.

RELEVANT_CLAIM will always contain a single number. Values in the source field like “1-4” will converted to 4 different records with the values 1, 2, 3 and 4.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2019 - First version

6.175 RESIDENCE_CTRY_CODE

Name: Country code of a person's residence
Also Known As: cc, country, territory
Description: The country of the residence - in contrast to the country of the correspondence addresses which is conventionally used.
Domain: 2 characters (A-Z), according to WIPO ST.3 or spaces
Exception: in case of bad data it may be any characters (e. g. "UK" is not ST.3 compliant, but should be "GB")
Default value: spaces
Source database
USPTO data of published applications and published grants

Source field name:
`<residence><country>`
Source sub-field identifier: n/a
Comments: n/a
This data is not available for applicants, only for inventors. Note that patent offices do not check the validity of the residence information.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version

6.176 ROLE

Name: The assignee's role according to the USPTO.
Also Known As: n/a
Description: Classification of an assignee as assigned by the USPTO.
Domain: 2 characters or empty,
The meaning of the values is as follows:
Note: A “1” in the first position identifies a partial owner.

01 or 11	Unassigned
02 or 12	United States company or corporation
03 or 13	Foreign company or corporation
04 or 14	United States individual
05 or 15	Foreign individual
06 or 16	U.S. Federal government
07 or 17	Foreign government
08 or 18	U.S. county government
09 or 19	U.S. state government

All other values are data errors and therefore they are not defined.

Default value: empty
Source database
USPTO data of published applications and published grants

Source field name:
`<assignee> <addressbook> <role>`
Source sub-field identifier: n/a
Comments: n/a
This data is not available for inventors, only for applicants.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version

6.177 SOURCE

Name: Name of the data source
Also Known As: n/a
Description: Identifies the source of the data in this record
Domain: 5 ASCII characters

DOCDB ➔ DOCDB, EPO's Bibliographic Database
EPREG ➔ EP Register
USPTO ➔ USPTO's Published Applications and Published Grants

Default value: n/a
Source database: n/a - Generated; value depends on data source
Source field name: n/a
Source sub-field identifier: n/a
Comments: The data source may be refined with the attribute SOURCE_VERSION.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version

6.178 SOURCE_VERSION

Name: Version of the data source
Also Known As: n/a
Description: Refines the attribute SOURCE
Domain: Up to 10 ASCII characters
If SOURCE = "DOCDB" ➔ SOURCE_VERSION is empty
If SOURCE = "EPREG" ➔ SOURCE_VERSION is empty
If SOURCE = "USPTO" ➔ SOURCE_VERSION is "BACKFILE", "4.2", "4.3", "4.4", …
Default value: n/a
Source database: n/a - Generated; value depends on data source
Source field name: n/a
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version

6.179 SPC_EXTENSION_DATE

Name: Extension date for SPC
Also Known As: n/a
Description: This is the last date of the validity of the granted Supplementary Protection Certificate.
Domain: Date
Default value: 9999-12-31
Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-details/spc/date-extension-granted

```
<legal-event providing-office="AT" date-added="20100805" date-previous-
exchange="20100812" sequence-number="6">
     <event-date>20100715</event-date>
     <event-date-effective>20100608</event-date-effective>
     <event-code>SZV</event-code>
     <event-details>
          <event-description event-description-type="original">SCHUTZZERTIFIKAT
VERLAENGERT</event-description>
          <event-description lang="en">SPC PROLONGED</event-description>
          <spc>
               <spc-number>SZ 31/96</spc-number>
               <date-filing>19961125</date-filing>
               <date-extension-granted>20110211</date-extension-granted>
               <text>PRODUCT NAME: ANASTROZOL; NAT. REG. NO/DATE: 301-21490
19960530; FIRST REG.: GB 12619/0106 19950811</text>
          </spc>
     </event-details>
</legal-event>
```

Comments
This attribute should not be confused with attribute SPC_PATENT_EXPIRY_DATE, which is typically *before* the SPC_EXTENSION_DATE.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.180 SPC_FILING_DATE

Name: SPC filing date
Also Known As: n/a
Description: Filing date of Supplementary Protection Certificate
Domain: Date
Default value: 9999-12-31
Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-details/spc/date-filing

```
<legal-event event-type="REG" providing-office="GB" date-added="20160716" date-
previous-exchange="20160716" sequence-number="57">
     <event-date>20160713</event-date>
     <event-code>CTFF</event-code>
     <event-details>
          <event-description event-description-type="original">SUPPLEMENTARY
PROTECTION CERTIFICATE FILED</event-description>
          <event-description lang="en">SUPPLEMENTARY PROTECTION CERTIFICATE
FILED</event-description>
          <spc>
               <spc-number>SPC/GB16/036</spc-number>
               <date-filing>20160615</date-filing>
               <text>PRODUCT NAME: TALIMOGENE LAHERPAREPVEC; REGISTERED: UK
EU/1/15/1064 20151218</text>
          </spc>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.181 SPC_NR

Name: SPC number
Also Known As: n/a
Description: Application and/or publication number of Supplementary Protection Certificate
Domain: Up to 40 ASCII characters
Default value: n/a
Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-details/spc/spc-number

```
<legal-event event-type="REG" providing-office="GB" date-added="20160716" date-
previous-exchange="20160716" sequence-number="57">
     <event-date>20160713</event-date>
     <event-code>CTFF</event-code>
     <event-details>
          <event-description event-description-type="original">SUPPLEMENTARY
PROTECTION CERTIFICATE FILED</event-description>
          <event-description lang="en">SUPPLEMENTARY PROTECTION CERTIFICATE
FILED</event-description>
          <spc>
               <spc-number>SPC/GB16/036</spc-number>
               <date-filing>20160615</date-filing>
               <text>PRODUCT NAME: TALIMOGENE LAHERPAREPVEC; REGISTERED: UK
EU/1/15/1064 20151218</text>
          </spc>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.182 SPC_PATENT_EXPIRY_DATE

Name: Expiry date of the patent on which the SPC is based
Also Known As: n/a
Description: The date the original underlying patent has expired
Domain: Date
Default value: 9999-12-31
Source database: INPADOC (EPO worldwide legal event database)
Source field name: /legal-status-document/legal-event/event-details/spc/date-expiry-of-patent

```
<legal-event providing-office="AT" date-added="20061121" date-previous-
exchange="20061123" sequence-number="3">
     <event-date>20061015</event-date>
     <event-code>EEZF</event-code>
     <event-details>
          <event-description event-description-type="original">ERTEILUNG EINES
SCHUTZZERTIFIKATES (E-SERIE)</event-description>
          <event-description lang="en">GRANT FOR A CERTIFICATE OF PROTECTION (E-
SERIES)</event-description>
          <spc>
               <spc-number>SZ   13/2002</spc-number>
               <date-filing>20020424</date-filing>
               <date-expiry-of-patent>20130622</date-expiry-of-patent>
               <text>SZ   13/2002, 20020424, EXPIRES:20130622</text>
          </spc>
     </event-details>
</legal-event>
```

Comments
This attribute should not be confused with attribute SPC_EXTENSION_DATE, which is typically *after* the SPC_PATENT_EXPIRY_DATE.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.183 SPC_TEXT

Name: Additional information for an SPC
Also Known As: n/a
Description: Additional information in free form text about a Supplementary Protection Certificate. It may contain product names, product registrations, product dates etc. as unstructured text.
Domain: up to 1 000 characters
Default value: empty

Source database: INPADOC (EPO worldwide legal event database)
Source field name:
/legal-status-document/legal-event/event-details/spc/text

```
<legal-event event-type="REG" providing-office="GB" date-added="20160716" date-
previous-exchange="20160716" sequence-number="57">
     <event-date>20160713</event-date>
     <event-code>CTFF</event-code>
     <event-details>
          <event-description event-description-type="original">SUPPLEMENTARY
PROTECTION CERTIFICATE FILED</event-description>
          <event-description lang="en">SUPPLEMENTARY PROTECTION CERTIFICATE
FILED</event-description>
          <spc>
               <spc-number>SPC/GB16/036</spc-number>
               <date-filing>20160615</date-filing>
               <text>PRODUCT NAME: TALIMOGENE LAHERPAREPVEC; REGISTERED: UK
EU/1/15/1064 20151218</text>
          </spc>
     </event-details>
</legal-event>
```

Comments
n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2017 - First version

6.184 ST3_NAME

Name: Name of country/territory or intergovernmental organisation
Also Known As: Country name
Description: Short English name of a state, other entity or intergovernmental organisation, as defined in WIPO standard WIPO ST.3 (plus minor extensions)
Domain: up to 100 characters
Default value: n/a
Source database: WIPO standard ST.3
Source field name: n/a
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version

6.185 STATE

Name: US state as part of the address
Also Known As: n/a
Description: Contains the US state as part of the address
Domain: Up to 2 ASCII characters or empty
Default value: empty string
Source database
USPTO data of published applications and published grants

Source field name:
`<addressbook> <address> <state>`
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version

6.186 STREET

Name: Street part of the address
Also Known As: n/a
Description: Contains the street part of the address
Domain: Up to 500 characters
Default value: empty string
Source database: USPTO data of published applications and published grants

Source field name:
`<addressbook> <address> <street>`
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version
M. Kracker - 01-10-2015 – Removed source “EP Register data”;
cf. attributes ADDRESS_1, ..., ADDRESS_5

6.187 TECH_REL_APPLN_ID

Name: Application identification of the technically related application
Also Known As: n/a
Description: Surrogate key based on the elements in the candidate primary key chosen
Domain: Number 1 … 999 999 999
Applications for which a technical relation had been found and for which no other relation is in existence. The relevant case is case # 5 from section 4.6 "Relation Types", using the `<priority-linkage-type>` value T.
Source database: DOCDB, PATSTAT
Source field name:

```
<application-reference is-representative="YES" data-format="docdb">
      <document-id>
            <country>US</country>
            <doc-number>44896706</doc-number>
            <kind>A</kind>
            <date>20060607</date>
      </document-id>
</application-reference>
.................
<language-of-publication>en</language-of-publication>
<priority-claims>
      <priority-claim sequence="1" data-format="docdb">
            <document-id>
                  <country>US</country>
                  <doc-number>44896706</doc-number>
                  <kind>A</kind>
                  <date>20060607</date>
            </document-id>
            <priority-active-indicator>Y</priority-active-indicator>
      </priority-claim>
      <priority-claim sequence="2" data-format="docdb">
            <document-id>
                  <country>US</country>
                  <doc-number>32859306</doc-number>
                  <kind>A</kind>
                  <date>20060110</date>
            </document-id>
            <priority-linkage-type>T</priority-linkage-type>
            <priority-active-indicator>N</priority-active-indicator>
      </priority-claim>
```

The corresponding application in PATSTAT is determined (via APPLN_AUTH, APPLN_NR and APPLN_KIND) and the value of APPLN_ID for this application is the TECH_REL_APPLN_ID

If there is no corresponding application in PATSTAT it should be created, see section 4.4 "Application replenishment".
Source sub-field identifier: n/a

Comments
This field defines the relationship between an application and a prior application (priority). If the priority-linkage-type = T, then the priority is a technically related priority.

Modification history
Author of update - Date of update - Explanation of update

R. Heijna - 20-07-2005 - Source field definition improved
J. Rollinson - 17-06-2009 - Changed source to DOCDB Exchange XML

6.188 TECHN_FIELD

Name: Name of a technology field
Also Known As: n/a
Description: English name of the technology field.
Domain: Up to 50 characters
Default value: n/a
Source database:
http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/xls/ipc_technology.xls

Source field name: See FIELD_EN of the above-mentioned Excel file

Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version

6.189 TECHN_FIELD_NR

Name: Number of a technology field
Also Known As: n/a
Description: Uniquely identifies a technology field. The number has little business meaning.
Domain: Number 1 … 35
Default value: n/a
Source database:
http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/xls/ipc_technology.xls

Source field name: See FIELD_NUMBER of the above-mentioned Excel file.
Source sub-field identifier: n/a

Comments: This is a classification according to *technology*. A classification according to *industries* is the NACE code which can be found in the tables TLS902_IPC_NACE2 and TLS229_APPLN_NACE2.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version
M. Kracker - 01-04-2015 –comment amended

6.190 TECHN_SECTOR

Name: Name of a technology sector
Also Known As: n/a
Description: The 35 technology fields are grouped in 5 technology sectors. This attribute contains the English name of the technology sector.
Domain: Up to 50 characters
Default value: n/a
Source database:
http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/xls/ipc_technology.xls

Source field name: See SECTOR_EN of the above-mentioned Excel file

Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2014 - First version

6.191 UNLESS_WITH_IPC

Name: IPC main group limiting the effect of attribute NOT_WITH_IPC
Also Known As: n/a
Description: Empty or first 8 characters of an IPC symbol according to WIPO ST.3.

Domain: Up to 8 ASCII characters; Example: `'A61K  8'`
Default value: empty
Source database: See Eurostat's paper described in section 5.28 "TLS902_IPC_NACE2: Mapping between IPC and industrial sectors".

Source field name: n/a
Source sub-field identifier: n/a
Comments: IPC main group which nullifies the effect of the attribute NOT_WITH_IPC column if it co-occurs with the symbol in the attribute IPC.
In the most cases this field is empty.

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2015 - First version

6.192 WEIGHT

Name: Weight of this association between the application and a classification according to an industry or a technical field.
Also Known As: n/a
Description: The higher the number, the stronger the relationship between an application and an industry / a technical field. The total of all weights of an application always equals 1.
Domain: Real number between 0 and 1
Default value: n/a
Source database: Computed from PATSTAT using reference table TLS902_IPC_NACE2 resp. TLS901_TECHN_FIELD_IPC.

Source field name: n/a
Source sub-field identifier: n/a
Comments:

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2015 - First version
M. Kracker - 01-10-2015 – Extended to be also applicable to Technical Fields

6.193 XP_NR

Name: XP number
Also Known As: n/a
Description: Non-patent literature (NPL) which have been handled by EPO examiners get an EPO internal accession number. Other NPLs, which have been used only by other offices, have the default number 0.
Domain: Number 0 … 950 000 000
Default value: 0
Source database: DOCDB, PATSTAT
Source field name:

```
<citation cited-phase="SEA" srep-office="EP" sequence="3">
      <nplcit num="1" npl-type="a" extracted-xp="000538241">
            <text>LI GONG ET AL:  "MULTICAST SECURITY AND ITS EXTENSION TO A
MOBILE ENVIRONMENT" WIRELESS NETWORKS, ACM, US, vol. 1, no. 3, 1 October 1995
(1995-10-01), pages 281-295, XP000538241 ISSN: 1022-0038</text>
      </nplcit>
      <category>X</category>
</citation>
```

In case of an NPL referring to a patent document, additionally there will be also a reference to the document-ID of the patent publication. This case should be treated the same as before, but the reference to the patent publication is kept as mentioned in the descriptions of table TLS212_CITATION and of attribute CITED_PAT_PUBLN_ID (see case c) in the table of the Business Rules in section 5.11 “TLS212_CITATION: Citation”.

Usage example showing a WPI abstract of a patent:

```
<citation cited-phase="SEA" srep-office="EP" sequence="3">
      <nplcit num="1" npl-type="a" extracted-xp="002594548">
            <text>DATABASE WPI Week 200214 Thomson Scientific, London, GB; AN
2002-099849 XP002594548 -&amp;amp; JP 2001 237065 A (TOPPAN PRINTING CO LTD) 31
August 2001 (2001-08-31)</text>
            <source-doc>
                  <document-id doc-id="289417544">
                        <country>JP</country>
                        <doc-number>2001237065</doc-number>
                        <kind>A</kind>
                  </document-id>
            </source-doc>
      </nplcit>
      <category>Y</category>
</citation>
```

Comments
Except for the default value 0, XP numbers uniquely identify an article, book, scientific paper, web page etc. However, because there often are variations in the way an article, book, etc. is referred too (see attribute NPL_BIBLIO), there may be multiple records in TLS214_NPL_PUBLN with the same XP_NR, but nevertheless referring to the same article, book, etc..

Modification history

Author of update - Date of update - Explanation of update
M. Kracker - 01-04-2021 – New attribute; Till 2020 Autumn Edition XP number was contained in the attribute NPL_PUBLN_ID

6.194 ZIP_CODE

Name: Zip code of the address
Also Known As: postal code, postcode
Description: Contains the zip code part of the address
Domain: Up to 30 characters
Default value: empty string
Source database
USPTO data of published applications and published grants

Source field name:
`<addressbook> <address> <postcode>`
Source sub-field identifier: n/a
Comments: n/a

Modification history
Author of update - Date of update - Explanation of update
M. Kracker - 01-10-2013 - First version

7 History of major changes to tables and attributes

October 2005	original
March 2006	Table tls213_npl_citn has been deleted. Table TLS212_pat_citn is now called tls212_citation (our citations data model needed reworking in December)
April 2006	TLS214_NPL_TEXT has been renamed to TLS214_NPL_PUBLN to comply with the Data Catalog. The indexes on TLS207_pers_appln changed from unique-primary on person_id & appln_id to non-unique on person_id and non-unique-clustering on appln_id
September 2006	No changes.
April 2007	No changes.
October 2007	New column IPC_CLASS_LEVEL in table tls209_appln_IPC to store the advanced or core indicator; new table TLS217_appln_I_CLS to store the applications classified in the Y01N In Computer Only EPO classification scheme for Nanotechnology.
April 2008	New table TLS218_docdb_fam - links applications which have exactly the same Paris Convention priorities in table TLS204_appln_prior; new web application to download parts of the data; TLS216 column renamed to parent_appln_id (used to be parent_appl_id)
September 2008	New table TLS219_inpadoc_fam links applications directly or indirectly - this corresponds to the extended INPADOC patent family in Espacenet or OPS web service.
April 2009	In table tls211_pat_publn, there is a new element publn_first_grant. This is a very tricky area - if you feel that you can improve on our rules, please discuss it with us via patstat@epo.org . The FAQs are now available on the internet forum; access to the forum is available via patstat@epo.org . In the area of citations, we no longer show the citations of patents in Non-Patent Literature (NPL) patent abstracts as separate citations. This means that before April 2009 if your programs were counting the citations for a patent, then your counts in April 2009 will be lower by the number of patents which were cited within patent abstracts. We no longer copy US inventor names into US assignee fields. The names and addresses from US Grants take precedence over those from US Published Applications. Referential integrity has been implemented for table tls207_pers_appln, to avoid applications having duplicate persons. Where possible (at the moment only for US data) , the person_name column in table tls206_person has been implemented as a concatenation of the last name, the first name and the middle names - separated by a comma.

	The separate files of person data, the TLS206_ASCII files, are intended only for users who wish to write special programs to process the name and address data. The last name, the first name and the middle names are stored in separate fields. Where the incoming data does not show the separation between the last first and middle names, then all are stored in the last-name field. For US data, the 'role' of the assignee is given. The web application which allows users to download a subset of the data is still active, but only for the edition of September 2008. The usage of the service has been monitored, and it has been decided not to extend this service beyond April 2009.
September 2009	Data coverage - PATSTAT has now increased data coverage for the Latin American countries: Guatemala (GT), Chile (CL), Ecuador (EC), Nicaragua (NI), Dominican Republic (DO), Panama (PA), Costa Rica (CR), Cuba (CU), Peru (PE) and El Salvador (SV). This will result in better family information as well. The web application which allowed users to download a subset of the data has been switched off.
April 2010	New table: Table TLS217_APPLN_I_CLS has been renamed TLS217_APPLN_ECLA and contains all ECLA codes and all ICO codes (including nanotech). This table covers ECLA (EPO Classifications scheme), ICO (InComputerOnly EPO Classification scheme,), IDT (Indeling der Techniek), and ECNO (ECLA symbols allocated by National Offices, not by EPO). The NANO-Technology symbols (Y01N) as provided in previous PATSTAT editions are part of the ICO scheme. This table contains extra columns that indicate the authority that assigned the code, the classification scheme and the symbol itself. You will need to keep this in mind for your scripts and queries. PATSTAT production process: the production of PATSTAT has been outsourced and the main data source is now the XML version of DOCDB. As a result of this change, the data quality has improved and a systematic user acceptance testing has been put in place. During this process we were able to eliminate about ten thousand duplicate "D2" applications. This new production process will guarantee a better synchronisation between DOCDB, PATSTAT and other patent information products. IPC related searches for documents published after 2006 allow now to find all the DOCDB simple family members consistently (in the previous editions you might have found only one or more but not all members of the family). This is due to the fact that, as (different) IPC classes can be present in DOCDB for all publication levels of an application, in PATSTAT these are now aggregated and de-duplicated at application level.
September 2010	PATSTAT production process: various adaptations have been introduced, the main ones being a) introduction of new citation sources in TLS212 (element 31 CITN_ORIGIN) and b) change of source to DOCDB XML for element PUBLN_FIRST_GRANT. The table TLS211_PAT_PUBLN contains the column PUBLN_FIRST_GRANT. If this has the value ‘1’, then that publication is the 'first grant'. In April 2010, the method for calculating this was based on the publication kind code representing a grant in each country, and then selecting the earliest publication. In September 2010 we use the 'public-availability' tag in the DOCDB XML product from the EPO.

	New table: Table TLS221_INPADOC_PRS containing INPADOC worldwide legal status data has been created and integrated into the PATSTAT database structure. However it has been produced on a test basis only, it will be available as of April 2011 edition but will have to be acquired separately.
April 2011	Table TLS201_APPLN: New permanent unique application identifier introduced in APPLN_ID. With the April 2011 edition, the DOCDB "doc-id" unique and stable identifier has been used to populate APPLN_ID instead of creating a PATSTAT-edition-specific surrogate key (*but not for the artificial applications in PATSTAT*). DOCDB attribute "doc-id" contains a stable and unique identifier that will allow for linking up a number of EPO raw data products through the application in a reliable way. This attribute will remain the same across PATSTAT editions and will always refer to the same combination of application authority, application number and application kind. Table TLS209_APPLN_IPC: IPC Core Level symbols are no longer maintained in WIPO ST8. Until September 2011, PATSTAT collected both the Advanced and Core sets of symbols for each application. The IPC Core symbols are now obsolete and have been eliminated from DOCDB, unless a publication had a Core symbol but no Advanced symbol. PATSTAT now shows the Advanced symbols, however Core symbols are still shown in those cases where no Advanced symbol is available but a Core symbol is still present in DOCDB.
October 2011	Creation of two additional tables TLS222 and TLS223 for JP and US national classifications, to reflect the additional national classification symbols now available in DOCDB XML. Addition of elements CITN_GENER_AUTH and CITED_APPLN_ID in citation table TLS212 to provide the International Search Authority (ISA) for PCT published application and to additionally provide details on cited applications (cited by the applicant). Addition of the number of claims for EP and US publications in table TLS211: the values are provided in the element PUBLN_CLAIMS. Improvements to existing data include adding 9 million abstracts in English language for Japanese published applications in table TLS203 To avoid confusion among table TLS210 and the newly created tables TLS222 and TLS223, all JP and US national classification symbols present in TLS210 have been removed.
April 2012	Table TLS214 NPL: the surrogate key has been replaced with XP number <refno> from DOCDB, this 9-digit number is used now as surrogate key. Element 57 PUBLN_CLAIMS: for the number of claims relating to EP B publications (granted patents) we have added the missing values for the years 1980-2005 in addition to the number of claims already available in October 2011 (2006 to date). For the US B publications (granted patents) we are now able to provide a timelier coverage (up to end 2011 in this edition).
October 2012	In table TLS201_APPLN, an additional rule has been implemented to further identify national applications originating from the PCT. This has raised the number of applications with an INTERNAT_APPLN_ID >0 from 4,850,479 to 5,319,404. In Table TLS212_CITATION, the domain for element 31 CITN_ORIGIN has changed: a new value PRS (for "PRe-Search" citations), to be attributed value 9 in PATSTAT, has been added.

	Remark: the ECLA classification scheme will be replaced by the new CPC - Cooperative Patent Classification on 01-01-2013. The next PATSTAT edition April 2013 will contain only CPC.
April 2013	• Table TLS224_APPLN_CPC has replaced TLS217_APPLN_ECLA. • Leading blanks in the attributes APPLN_NR and PUBLN_NR have been removed. • The APPLN_ID ranges for artificial applications and the PUBLN_ID ranges for artificial publications are now fixed and do not depend anymore on the number of applications / publications. • Artificial applications which do not have a publication (= their APPLN_ID is > 930 000 000) are not assigned to any INPADOC family
October 2013	• EP Register is used (again) instead of EP Bulletin as a data source for names and address of applicants and inventors of EP patents. • New table TLS226_PERSON_ORIG which replaces the previous file TLS206_PERSON_ASCII. It contains the unmodified name and address data for all persons in PATSTAT • New table TLS227_PERS_PUBLN which links applicants and inventors to publications. • Attributes PERSON_ID and PERSON_ORIG_ID are from now on stable • Freeform name strings from the USPTO data source are constructed more consistently with other freeform names • De-duplication rules for table TLS206_PERSONS are changed • New attribute TLS201_APPLN.APPLN_EPODOC_NR to easily connect to Espacenet etc. • Additional values for TLS212_CITATION.CITN_ORIGIN and reordering of columns • Unused columns of TLS221_INPADOC_PRS removed: L514EP, L521EP,L526EP, L527EP • Change in computation of TLS201_APPLN.IPR_TYPE
2014 Spring Edition	• New tables TLS801_COUNTRY and TLS802_LEGAL EVENT_CODE • Table TLS901_TECHN_FIELD_IPC replaces table TECHN_FIELD_IPC • New attribute LEC_ID in table TLS221_INPADOC_PRS • New default values for dummy applications (APPLN_ID = 0) and dummy publications (PAT_PUBLN_ID = 0) • Attribute L519EP has extended string length • Attribute L520EP is now numeric • Attribute TECHN_FIELD in table TLS209_APPLN_IPC renamed to TECHN_FIELD_NR • Value “P” (provisional application) defined for attribute APPLN_KIND
2014 Autumn Edition	• Double quotes (“) are replaced by single quotes (‘) • PUBLN_ID is now a stable attribute • IPC classification symbol which are classified on subclass level only are now included (IPC_CLASS_LEVEL = ‘S’)

	• Deduplication rules when merging IPCs from different publications into their application has been slightly adapted. Same rules apply also for CPC classification symbols.
2015 Spring Edition	• Several tables and attributes which were available only in PATSTAT Online are now also in PATSTAT Raw Data: details see below. • New table TLS906_PERSON has been added. It includes all data from TLS206_PERSON, plus all data about harmonised names previously only available for PATSTAT Online. It is advised that users replace the original table TLS206_PERSON by this new extended table TLS906_PERSON. • Table TLS208_DOC_STD_NMS has been removed and its content integrated into TLS206_PERSON. • Table INDUSTRY_IPC has been replaced by table TLS902_IPC_NACE, which is now also available in PATSTAT Raw Data. The new table represents the new IPC – NACE concordance table published by Eurostat in 2014. • New table TLS229_APPLN_NACE2 has been added. • Table DOCDB_FAMILY_CITATION has been renamed to TLS228_DOCDB_FAM_CITN and made available to PATSTAT Raw Data. Its attributes have been re-ordered and partly renamed. • New column ISO_ALPHA3 has been added to table TLS801_COUNTRY. • In table TLS201_APPLN the name of attribute NB_CITATIONS has been changed to NB_CITING_DOCDB_FAM. • The order of attributes has changed in table TLS901_TECHN_FIELD_IPC. • The attribute DOC_STD_NAME_ID is no longer a stable ID. • Duplicates in persons have been removed. • Some computed attributes of PATSTAT Online are regarded as deprecated and may be removed in future editions: In table TLS201_APPLN: - APPLN_FILING_YEAR_MONTH - APPLN_FILING_YEAR - PRIOR_EARLIEST_YEAR_MONTH - PRIOR_EARLIEST_YEAR - PUBLN_EARLIEST_YEAR_MONTH - PUBLN_EARLIEST_YEAR - DOCDB_FAMILY_ID In table TLS209_APPLN_IPC: - IPC_SUBCLASS_SYMBOL In table TLS224_APPLN_CPC: - CPC_MAINGROUP SYMBOL
2015 Autumn Edition	• Removed differences between the data models of PATSTAT Raw Data and PATSTAT Online, by removing these pre-computed and redundant attributes: - APPLN_FILING_YEAR_MONTH from TLS201_APPLN - PRIOR_EARLIEST_YEAR_MONTH from TLS201_APPLN - PUBLN_EARLIEST_YEAR_MONTH from TLS201_APPLN - PUBLN_EARLIEST_REF from TLS201_APPLN - IPC_SUBCLASS_SYMBOL from TLS209_APPLN_IPC - TECHN_FIELD_NR from TLS209_APPLN_IPC - CPC_MAINGROUP_SYMBOL from TLS224_APPLN_CPC. • Re-ordered attributes in table TLS201_APPLN and added attribute EARLIEST_FILING_ID • Renamed these attributes of TLS201_APPLN: - PRIOR_EARLIEST_DATE to EARLIEST_FILING_DATE

	- PRIOR_EARLIEST_YEAR to EARLIEST_FILING_YEAR - PUBLN_EARLIEST_DATE to EARLIEST_PUBLN_DATE - PUBLN_EARLIEST_YEAR to EARLIEST_PUBLN_YEAR • Moved attribute APPLN_TITLE_LG from TLS201_APPLN to TLS202_APPLN_TITLE • Moved attribute APPLN_ABSTRACT_LG from TLS201_APPLN to TLS203_APPLN_ABSTRACT • Table TLS212_CITATION: In attribute CITN_ORIGIN code “115” has been renamed to “TPO” (Third Party Observation” • Tables TLS218_DOCDB_FAM and TLS219_INPADOC_FAM have been integrated into table TLS201_APPLN • TLS226_PERSON_ORIG: 5 new attributes for 5 address lines for addresses of EP applications. • New table TLS230_APPLN_TECHN_FIELD
2015 Autumn Edition - Amended	• Extended the domains for attributes DOCDB_STD_NAME, CITN_ID, PAT_CITN_SEQ_NR and NPL_CITN_SEQ_NR • New processing rules for attributes of table TLS212_CITATION, to adapt for changes in DOCDB
2016 Spring	• Table TLS201_APPLN: New attributes APPLN_NR_ORIGINAL, INT_PHASE, REG_PHASE and NAT_PHASE. • Tables 206_PERSON and 906_PERSON: Attributes for EEE-PPAT names have been renamed from "HRM_..." to "PSN_..." (PATSTAT Standardised Name) and are now also available in TLS206_PERSON. Attribute HRM_L1 has been removed. The length of all person attributes has been unified to 500 characters. • Table TLS211_PAT_PUBLN: New attribute PUBLN_NR_ORIGINAL • Table TLS212_CITATION: Attribute NPL_PUBLN_ID has been renamed to CITED_NPL_PUBLN_ID • Table TLS214_NPL_PUBLN: New attribute NPL_TYPE. NPL references which contain in their biblio text something like "none” or "See also references ..." are removed. • TLS226_PERSON_ORIG: Attributes ADDRESS_1 - ADDRESS_3 are now also used for the USPTO data source. • TLS229_APPLN_NACE2: Due to an update by Eurostat the logic has slightly changed.
2016 Autumn	• Line breaks are replaced by " \n" in text attributes. • All artificial applications now also belong to a DOCDB family and an INPADOC family. They will have no other family member. See attributes DOCDB_FAMILY_ID and INPADOC_FAMILY_ID. • Table TLS205_TECH_REL now also contain inverse relations, because technical relations are by definition symmetric. • Regionalisation information has been added: - new attributes NUTS and NUTS_LEVEL in table TLS906_PERSON - new reference table TLS904_NUTS
2017 Spring	• Table TLS221_INPADOC_PRS has been replaced by table TLS231_INPADOC_LEGAL_EVENT • Table TLS802_LEGAL_EVENT_CODE has been replaced by table TLS803_LEGAL_EVENT_CODE

	• 18 attributes have been added to table TLS214_NPL_PUBLN. Note that – depending on the type of the Non-Patent Literature and the attribute – attributes may be populated sparsely. • In attribute CONTINENT the value for “America” has been replaced by “North America” resp. “South America”

2017 Autumn	• Table TLS214_NPL_PUBLN
	  • Some attributes are now populated for more NPL types.
	  • Attribute ONLINE_CLASSIFICATION may hold more than one Derwent class.
	  • Attribute ONLINE_AVAILABILITY can now hold up to 500 characters.
	  • Attribute NPL_AUTHOR now can hold up to 1 000 characters.
	• Table TLS231_INPADOC_LEGAL_EVENT:
	  • New attribute EVENT_FILING_DATE

2018 Spring	• Table TLS201_APPLN
	  • Attribute APPLN_AUTH has changed its meaning. For international applications it now contains “WO” and not the receiving office anymore. Please adapt your existing script.
	  • Attribute RECEIVING_OFFICE has been added, which is also part of the alternate key.
	• Citations (tables TLS212_CITATION and TLS215_CITN_CATEG):
	  • For Euro-PCT publications (i.e. international applications in the EP regional phase) the citations of the international publications have been added to the citations of the EP publication.
	  Reason: These citations are not re-published by the EPO, although they must be considered when performing citation analysis.
	  A new attribute CITN_REPLENISHED contains the information whether a citation has been replenished and from which publication the citation originates.
	• Table TLS231_INPADOC_LEGAL_EVENT
	  • New attribute EVENT_ID, which is a stable identifier for legal events.
	• Attribute “CITN_CATEG”: “&” is now a valid value.
	• NUTS territorial codes
	  • The reference table TLS904_NUTS has been restructured. It contains now NUTS codes from levels 1 to 3.
	  • NUTS codes (see attribute NUTS) is now conformant to the official NUTS codes in the reference table.

2018 Autumn	• Table TLS803_LEGAL_EVENT_CODE
	  • Change of name and content of attributes for legal event categories which are now based on WIPO ST.27
	  • Attribute EVENT_IMPACT is deprecated.
	• DOCDB specific language codes “bs”, “hr” and “me” are added to domains of attributes APPLN_ABSTRACT_LG and APPLN_TITLE_LG
	• Some flag attributes changed their domain from 0/1 to Y/N:
	  • Attribute GRANTED in table TLS201_APPLN
	  • Attribute PUBLN_FIRST_GRANT in table TLS211_PAT_PUBLN
	• The data source for computing the value of attribute GRANTED of table TLS201_APPLN has been improved
	• Attribute NPL_BIBLIO in TLS214_NPL_PUBLN: The domain has been extended to allow longer strings.

2019 Spring	• New attribute PERSON_NAME_ORIG_LG in tables
	  • TLS206_PERSON
	  • TLS906_PERSON

	• TLS226_PERSON_ORIG Because the new attribute is also part of the Primary Key, persons which have been represented in the past by a single record in a person table are now represented by multiple records in case their original language name differs. • Table TLS215_CITN_CATEG: • New attribute RELEVANT_CLAIM • Attribute CITN_CATEG may hold multiple citation categories if they are all related to the same relevant claim • Attribute APPLN_NR_EPODOC in table TLS201_APPLN is deprecated.
2019 Autumn	• Table TLS906_PERSON has been removed. Its content has been merged into table TLS206_PERSON. • Table TLS207_PERS_APPLN: Now this table links the applicants and inventors of the most recent publication which contains Latin person names to an application. Publications which contain only persons with non-Latin names (e.g. with Chinese characters) are not considered here anymore. • Table TLS201_APPLN: Values for NB_APPLICANTS and NB_INVENTORS have been slightly redefined. • CPC_GENER_AUTH is also populated for CPC_SCHEME="CPC" • NPL_PUBLN_ID can have more than 9 digits
2020 Spring	• Due to a change practise of the EPO and USPTO, CPC classification symbols are not assigned anymore to applications, but to DOCDB families. The new table TLS225_DOCDB_FAM_CPC reflects this. It also contains information previously not available. The existing table TLS224_APPLN_CPC with its most relevant attributes is redundantly kept for downward compatibility. • TableTLS216_APPLN_CONT now contains also changes of IPR type (changes from patents to utility model to vice versa). • The mapping of applicants, inventors and assignees of the USPTO data source to applicants and inventors in PATSTAT has been improved. • In table TLS801_COUNTRY the attribute STATE_INDICATOR has been renamed to ORGANISATION_FLAG and its domain changed accordingly. • Table TLS906_PERSON has been removed. Its content has been merged into table TLS206_PERSON.
2020 Autumn	• Euro-PCTs are replenished with abstracts from PCTs • Some NUTS values in TLS206_PERSON are provided by OECD's REGPAT database, January 2020. These enhanced records have a NUTS_LEVEL with value 4.
2021 Spring	• Change in table TLS214_NPL_PUBL: - Duplicates have been consolidated - Key attribute NPL_PUBLN_ID is not a number anymore, but the MD5 hash of attribute NPL_BIBLIO - XP-number – if it exists – is now in a separate attribute XP_NR (was in NPL_PUBLN_ID before) • Domain of attribute CITED_NPL_PUBLN_ID in table TLS212_CITATION has been changed to be consistent with attribute NPL_PUBLN_ID of table TLS214_NPL_PUBLN.

	• In table TLS231_INPADOC_LEGAL_EVENT: New legal event codes for EPO's validation states have been introduced, notably VSFP (annual fee payment in a validation state) and VS25 (lapse in validation state).
2021 Autumn	• No changes

8 Known deficiencies

Data coverage issues are not described in this section. Data coverage information can be found in https://public.tableau.com/profile/patstat.support#!/vizhome/CoverageofPATSTAT2021SpringEdition/CoveragePATSTATGlobal and https://www.epo.org/searching-for-patents/data/coverage/weekly.html .

- **TLS206_PERSON: DOCDB standardised names**
  Some DOCDB standardised names are wrongly assigned to persons of US patents, because the sequence of persons in the USPTO data source and that in DOCDB sometimes do not match correctly.
  There is no known fix. When working with US patent applicants or inventors, you should avoid using the DOCDB standardised name. Instead, you might consider other harmonised names available in table TLS206_PERSON.

- **TLS206_PERSON: NUTS regions**
  Addresses of persons which have been added in the 2019 Spring Edition or later have not been regionalised. That means that for these persons the attributes NUTS and NUTS_LEVEL contain default values only.
  To rectify this situation, some NUTS values in TLS206_PERSON are enhanced by OECD's REGPAT database, January 2020. These enhanced records have a NUTS_LEVEL with value 4.

- **TLS225_DOCDB_FAM_CPC: CPC_POSITION**

  Only one CPC symbol per family and generating authority should have CPC_POSITION equal to 'F'. Please note that currently there are cases of multiple 'F' assignments, due to some technical issues with the DOCDB backfile extraction (CPC_POSITION is based on DOCDB symbol-position).