The REG118_PREV_FILED_APPLN Table

Welcome to a comprehensive exploration of one of the key tables in the PATSTAT database: the `REG118_PREV_FILED_APPLN` table. This table plays a pivotal role in tracking previously filed patent applications, providing insights into the continuity and progression of patent families across jurisdictions and time.

The table includes essential attributes such as `ID` (a unique numerical identifier for an application), `APPLN_AUTH` (the authority where the application was filed), `APPLN_NR` (the application number as issued by the filing authority), and `APPLN_DATE` (the date of filing). It also features publication-related information like `BULLETIN_YEAR` and `BULLETIN_NR`, indicating the year and calendar week of the EPO Bulletin in which the relevant actions were published.

By linking `REG118_PREV_FILED_APPLN` with other tables, such as `REG101_APPLN` (which contains information on current applications) or `REG201_PROC_STEP` (which details procedural steps), we can gain a deeper understanding of how applications evolve and are connected within patent families. This linkage helps map the lifecycle of patent applications, from filing through to publication and beyond.

The `APPLN_AUTH` and `APPLN_DATE` fields enable geographic and temporal analysis of patent filings. For instance, researchers can identify trends in filing behaviors across regions or examine the impact of filing dates on the grant process. The `BULLETIN_YEAR` and `BULLETIN_NR` attributes provide a direct connection to EPO Bulletin publications, enabling users to pinpoint the timing and context of key events.

In [1]:
```
from epo.tipdata.patstat import PatstatClient
from epo.tipdata.patstat.database.models import REG118_PREV_FILED
_APPLN, REG101_APPLN
from sqlalchemy import select, func, case, select, and_

patstat = PatstatClient(env='PROD')

db = patstat.orm()
```

In [2]:
```
q = db.query(
    REG118_PREV_FILED_APPLN.id,
    REG118_PREV_FILED_APPLN.bulletin_year,
    REG118_PREV_FILED_APPLN.bulletin_nr,
    REG118_PREV_FILED_APPLN.appln_auth,
    REG118_PREV_FILED_APPLN.appln_nr,
    REG118_PREV_FILED_APPLN.appln_date,
    REG101_APPLN.appln_filing_date
).join(
    REG101_APPLN, REG118_PREV_FILED_APPLN.id == REG101_APPLN.id
# Join on ID
)

res = patstat.df(q)
res
```

Out[2]:

	id	bulletin_year	bulletin_nr	appln_auth	appln_nr	appln_date	appln_fil
0	10002862	2010	24	EP	20080004010	2005-08-24	20
1	15001476	2016	2	EP	20140168516	2014-05-15	20
2	20171899	2020	51	EP	20120712105	2012-03-29	20
3	22155477	2022	36	WO	2019EP60221	2019-04-19	20
4	18161209	2018	51	WO	2009NL50803	2009-12-24	20
...	...	...	...	...	...	...	
6658	21200427	2023	2	CN	202110773990	2021-07-08	20
6659	10183059	2011	17	EP	20090013894	1996-04-24	19
6660	16174405	2017	4	WO	2012EP73255	2012-11-21	20
6661	13175921	2013	46	EP	20080749598	2008-04-17	20
6662	18155648	2018	45	WO	2013IL50607	2013-07-17	20

6663 rows × 7 columns

Key Fields in the `REG118_PREV_FILED_APPLN` Table

ID (Primary Key)

The ID field serves as a technical identifier that uniquely connects patent applications across various tables. The ID is essential for linking application data with other related information.

ID values are derived systematically, ensuring that a specific application maintains the same ID in all PATSTAT editions. For European Patent (EP) applications, the ID is derived from the XML attribute id by removing the prefix "EP," suffix "P," and any leading zeros.

BULLETIN_YEAR

In the PATSTAT database, the `BULLETIN_YEAR` field captures the year when an action or event related to a patent application was published in the EPO Bulletin.

The `BULLETIN_YEAR` is a 4-digit numeric field (formatted as YYYY), with a default value of 0 to indicate cases where no bulletin publication is known. For entries where publication in the EPO Bulletin is confirmed, `BULLETIN_YEAR` reflects the corresponding year of publication. It is used in conjunction with `BULLETIN_NR`, which specifies the bulletin issue number.

BULLETIN_NR

The `BULLETIN_NR` attribute represents the issue number of the EPO Bulletin in which a specific action has been published. This number indicates the calendar week during which the Bulletin was released. It serves as a reference for identifying the exact edition of the EPO Bulletin where actions such as patent grants, publications, or other significant events are announced.

If the action was not published in the Bulletin or if the information is unknown, the default value of 0 is used for the `BULLETIN_NR`, which corresponds to the absence of a known bulletin number. This value is only used when the associated `BULLETIN_YEAR` is also set to 0.

In [20]:
```
q = db.query(
    REG118_PREV_FILED_APPLN.id,
    REG118_PREV_FILED_APPLN.bulletin_year,
    REG118_PREV_FILED_APPLN.bulletin_nr,
    REG118_PREV_FILED_APPLN.appln_auth,
    REG118_PREV_FILED_APPLN.appln_nr,
    REG118_PREV_FILED_APPLN.appln_date,
).filter(
    REG118_PREV_FILED_APPLN.bulletin_year == 2014
).order_by(
    REG118_PREV_FILED_APPLN.bulletin_nr
)

res = patstat.df(q)
res
```

Out[20]:

	id	bulletin_year	bulletin_nr	appln_auth	appln_nr	appln_date
0	13174977	2014	1	EP	20060253236	2006-06-22
1	14002925	2014	1	WO	2006EP04278	2006-05-08
2	13162439	2014	1	WO	2007CN02967	2007-10-16
3	13166384	2014	1	WO	2007US82093	2007-10-22
4	14172757	2014	1	EP	20080000925	2000-09-28
...	...	...	...	...	...	...
317	14001709	2014	51	DE	20131009998	2013-06-14
318	14170945	2014	51	US	201313914178	2013-06-10
319	14169054	2014	52	EP	20070766873	2007-07-05
320	14177377	2014	52	WO	2009EP61732	2009-09-10
321	14181462	2014	52	EP	20120165218	2008-05-19

322 rows × 6 columns

APPLN_AUTH

The `APPLN_AUTH` attribute denotes the authority or office where a patent application has been filed. For other patent applications, the `APPLN_AUTH` field may represent different authorities, depending on the jurisdiction where the application was submitted. The value is typically a two-character code, assigned according to the WIPO ST.3 standard, which defines country and regional office codes.

This attribute is essential for identifying the filing authority behind each patent application, providing insights into the geographical and institutional origins of patent filings. The default value is not applicable, as the `APPLN_AUTH` field is always populated with the respective authority code, ensuring clarity on where an application has been officially lodged.

In [23]:
```
q = db.query(
    REG118_PREV_FILED_APPLN.appln_auth,
    func.count(REG118_PREV_FILED_APPLN.appln_nr).label('application_count')
).group_by(
    REG118_PREV_FILED_APPLN.appln_auth
).order_by(
    func.count(REG118_PREV_FILED_APPLN.appln_nr).desc()  # Order by count in descending order
)


res = patstat.df(q)
res
```

Out[23]:

	appln_auth	application_count
0	WO	3603
1	EP	2337
2	DE	256
3	US	195
4	PL	77
5	FR	28
6	CN	27
7	ES	20
8	GB	15
9	JP	15
10	CZ	10
11	IT	9
12	CH	9
13	NL	8
14	SE	7

15	AT	5
16	SI	5
17	IL	5
18	GR	4
19	TW	4
20	KR	4
21	TR	4
22	RO	3
23	BE	2
24	FI	2
25	HR	2
26	MX	1
27	AU	1
28	ME	1
29	UA	1
30	NO	1
31	HU	1
32	SG	1

APPLN_NR

The APPLN_NR attribute represents the application number assigned by the application authority. This unique identifier is issued by the relevant patent office and is used to track and reference a specific patent application.

There is no default value for this attribute, as each application is assigned a unique number upon filing, and this number must be provided for each record.

The domain for this attribute consists of a string of 8 digits. Leading zeros are significant.

APPLN_DATE

The `APPLN_DATE` attribute represents the date when a patent application was filed. In the `REG118_PREV_FILED_APPLN` table, this field captures the exact filing date for each application, providing temporal information for tracking the progress of patent filings over time.

The default value for this attribute is set to 9999-12-31, which serves as a placeholder to indicate an unknown or unspecified filing date.

In [ ]: