

# The Citation between DOCDB Families Table (TLS228\_DOCDB\_FAM\_CITN)

Welcome to the **Citation between DOCDB Families Table**, coded as TLS228\_DOCDB\_FAM\_CITN. 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.

In this table, both cited publications (typical case) and cited applications (also possible) are considered.

```
In [6]: from epo.tipdata.patstat import PatstatClient

# Initialize the PATSTAT client
patstat = PatstatClient(env='PROD')

# Access ORM
db = patstat.orm()

# Importing the as models
from epo.tipdata.patstat.database.models import TLS228_DOCDB_FAM_CITN
```

## DOCDB\_FAMILY\_ID

This attribute provides a link to table TLS201 and it is the identifier of a DOCDB simple family.

```
In [7]: # Import table TLS201
from epo.tipdata.patstat.database.models import TLS201_APPLN

show_join = db.query(
    TLS201_APPLN.docdb_family_id,
    TLS228_DOCDB_FAM_CITN.cited_docdb_family_id
).join(
    TLS201_APPLN, TLS228_DOCDB_FAM_CITN.docdb_family_id == TLS201_APPLN.docdb_family_id
).limit(1000)

show_join_df = patstat.df(show_join)
show_join_df
```

Out[7]:

	docdb_family_id	cited_docdb_family_id
0	71090642	41033455
1	41011909	19062236
2	41011909	19062236
3	41011909	19062236
4	38777173	37295458
...	...	...
995	3800217	13063532
996	3800217	13063532
997	3800217	13063532
998	3800217	13063532
999	3800217	13063532

1000 rows × 2 columns

## CITED\_DOCDB\_FAMILY\_ID

ID of the cited DOCDB simple family. Uniquely identifies the cited family.

With this attribute, we can for example retrieve the 5 most cited DOCDB family among the ones having applications filed at the EPO.

```
In [9]: from sqlalchemy import func

most_cited_fams = db.query(
    TLS228_DOCDB_FAM_CITN.cited_docdb_family_id,
    func.count(TLS201_APPLN.docdb_family_id).label('Total citations')
).join(
    TLS228_DOCDB_FAM_CITN, TLS201_APPLN.docdb_family_id == TLS228_DOCDB_FAM_CITN.docdb_family_id
).filter(
    TLS201_APPLN.appln_auth == 'EP'
).group_by(
    TLS228_DOCDB_FAM_CITN.cited_docdb_family_id
).order_by(
    func.count(TLS201_APPLN.docdb_family_id).desc()
).limit(5)

most_cited_fams_df = patstat.df(most_cited_fams)
most_cited_fams_df
```

Out[9]:

	cited_docdb_family_id	Total citations
0	23920101	9079
1	27109640	7219
2	24543134	6080
3	27124781	5708
4	27399321	4514

In [ ]: