

Risis Patent Database

RISIS Training Sessions

Patricia Laurens

LISIS, CNRS, UGE



RISIS PATENT DATABASE (RPD)

- The RISIS Patent database is designed for the analysis of technological knowledge creation, using patent as a proxy.
- It derives from the EPO PATSTAT (Version April 2017; Version Spring 2021 to come).
- It contains bibliographical patent data from leading industrialised and developing countries.
- It focuses on ‘priority patents’, the very first patent application for new invention, that represent the creation of new knowledge. It includes priority patents labelled as “artificials” in Patsat.
- RPD offers a series of improvements (filling of missing data, additions of new information, calculations of indicators).

RISIS PATENT DATABASE (RPD)

RPD provides essential patent information for analysing the geography and the actors of the knowledge production, the content and the value of the knowledge production.

GEOGRAPHY – ACTORS - TECHNOLOGY

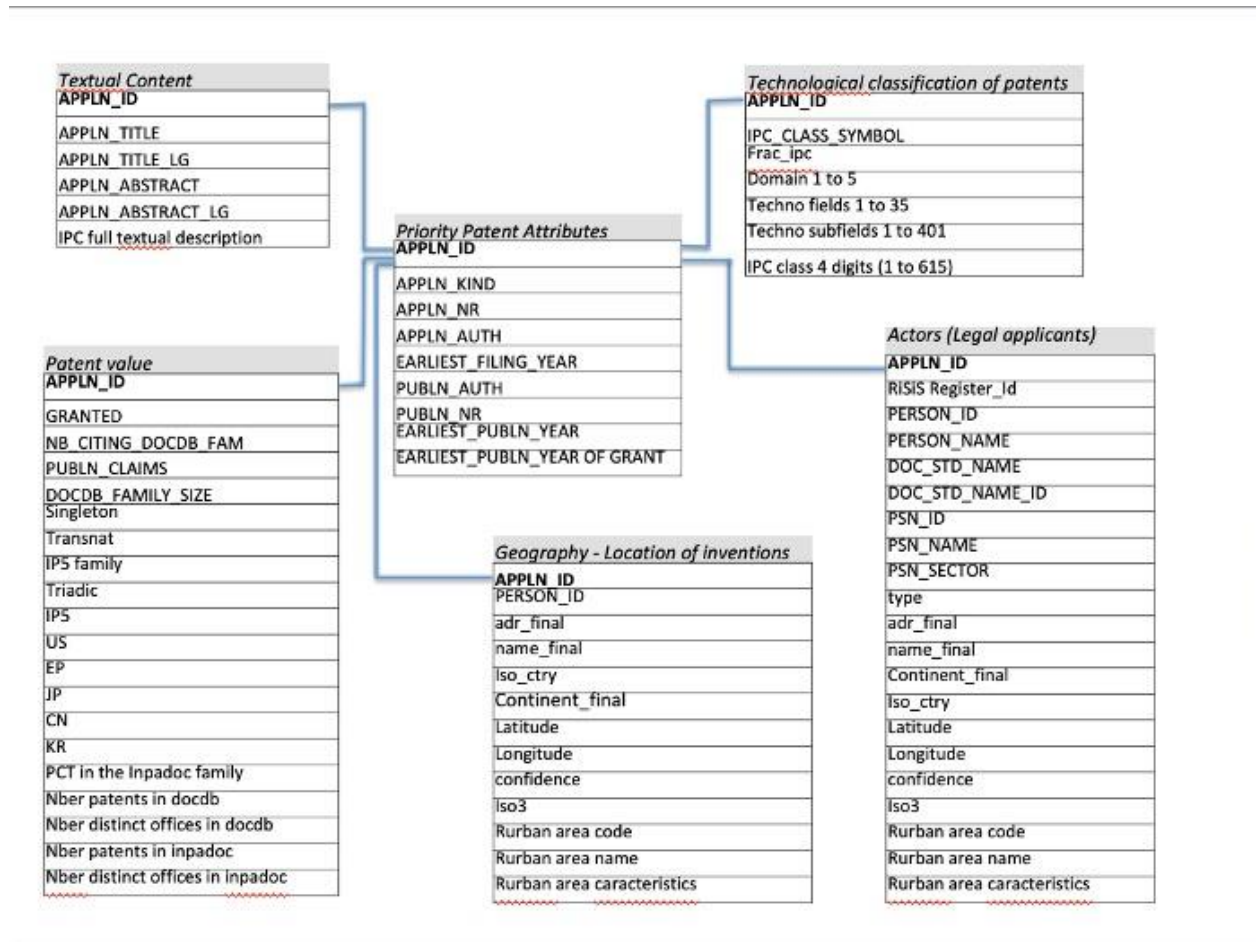
- **Legal attributes** of the priority patent applications (data from Patstat)
- **Content of invention**, using textual pieces of information (Patstat (title,abstract) +IPC description)
- **Technological content** using the standard technology classification (IPC (CPC) from Patstat + IPC aggregation: domains, fields, subfields, fractional counts)
- **Geographical location of inventive activities** based on inventors' addresses (Patstat + latitude/longitude, NUTS2, URA allocation, fractional counts)
- **Legal organisations** that apply for patents (Pastat + sector allocation (public organisations/large firms/startup/FGMF, harmonisation of applicants' names/link with ORGREG/ETER and FIRMREG)
- Characteristics linked to the **'value' of the patent** (Patstat + info on geographical extension: singleton/transnational/application in IP5 offices, triadic, **2IP5 families**)

RISIS PATENT DATABASE (RPD)

- Filling of information in RPD use internal or external sources
 - Filling information in Artificial patents (priority patents without information in Patstat) -13% of priority patents
 - Completing addresses of inventors: The process allows to end in RISIS Patent with 75% of inventors with an address (to be compared with 10% in the initial raw PATSTAT data)
 - Completing names of applicants (artificial patents) – Matching with actors' names
- Addition of new types of data
 - Classification of applicant type, harmonisation of applicants' names, link with ORGREG and FIRMREG repositories
 - Geocoding and allocation to geographical areas: 67.4% of the addresses are geocoded. The share of geocoded addresses varies according to countries. It exceeds 80% in most of the western countries.

RISIS PATENT DATABASE (RPD)

RPD is a relational database using a Mysql querying environment.



RISIS PATENT DATABASE (RPD)

- 16,406,977 priority patents of inventions were applied for from 2000 to 2015. They include 39,761,538 inventors.
- Today, RPD is accessible to researchers with a research project – Applications via the RISIS portal (<https://rcf.risis2.eu/datasets>)
- The selection of the criteria to define the patent data for a given research project is a key issue:
 - How to delineate a technological area? (IPC codes/Lexical/Actors)
 - Which type of priority patents ? Large coverage with all priority ? (office bias) vs a selection of patents with a threshold value (2IP5 family priority patents)

RISIS PATENT DATABASE (RPD)

Selection of the priority patents –Value of patents

Priority patents applied in several countries (higher economical value, reduce national bias)

Patents applied in main economical zones: US, EU, JP, CN, KR

Type of patents	Share
Transnat priority patents	20,2%
2IP5 family priority patents	16,5%
PCT family priority patents	18,5%
Triadic priority patents	5,5%
Priority patents	100,0%

2IP5 priority patents: patents applied in a least 2 IP5 offices as a marker for internationally valuable technology output of R&D → International Patents

- 2010-2014: 716 160 Intl. geocoded priority patents
- 2000-2004: 542 492 Intl. geocoded priority patents

Corporate Invention board (CIB2)

CIB patent database: a database of priority patents applied worldwide from 2000 to 2015 by the 4000 largest techno innovative companies worldwide – aggregated perimeter of the firm (share >50.01%)

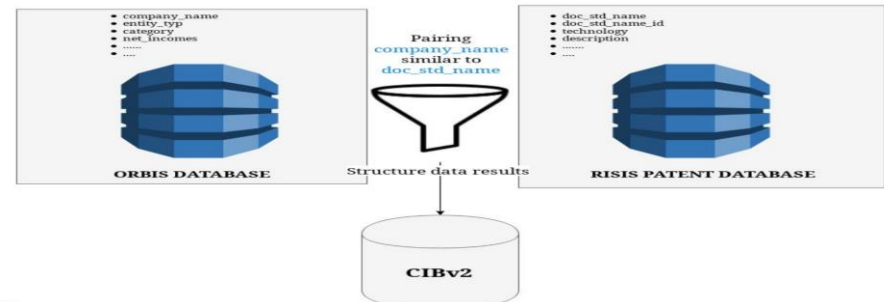
It was developed using:

Patent data (**RISIS patent database**)

Data on firms with R&D activities activities:

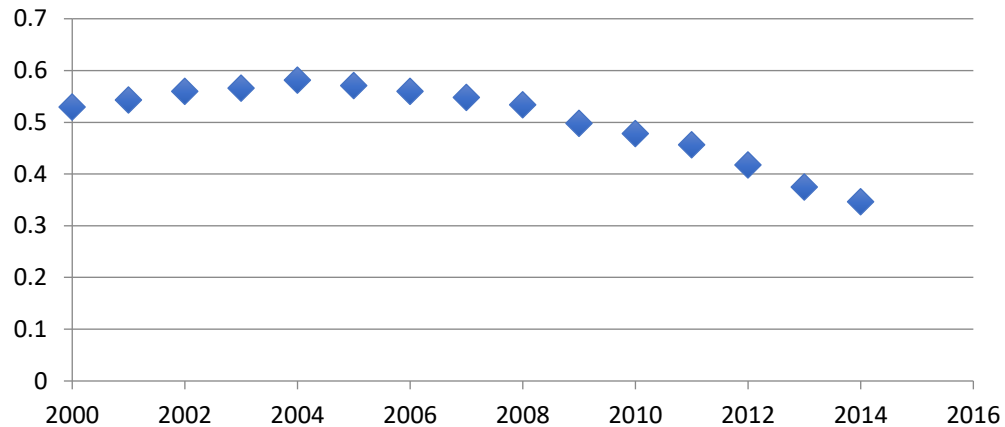
- R&D investments (**R&D Industrial Scoreboards**)
- Patenting activities (**WIPO top patenting firms**)
- **Orbis database** (bureau Van Dijk) : parent companies and the list of their subsidiaries

Matching of applicant & ORBIS firm names (more than 300000)
in 13 millions priority patents



CIB2: Overview

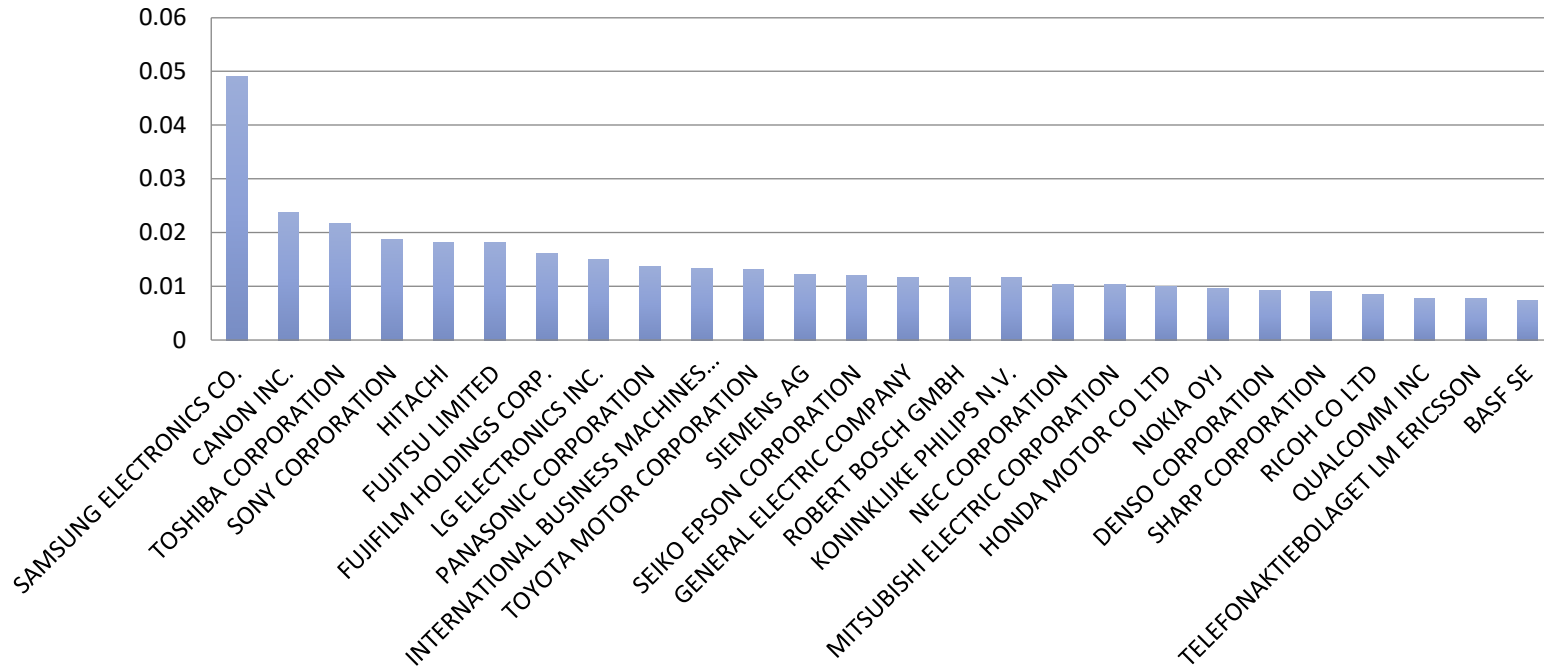
Share of CIB2 patents in Risis Patent Database



- **6 173 089** priority patents
- **1 733 439** IP5 family patents
- **676 916** triadic patents
- **286 132** IP5 patents

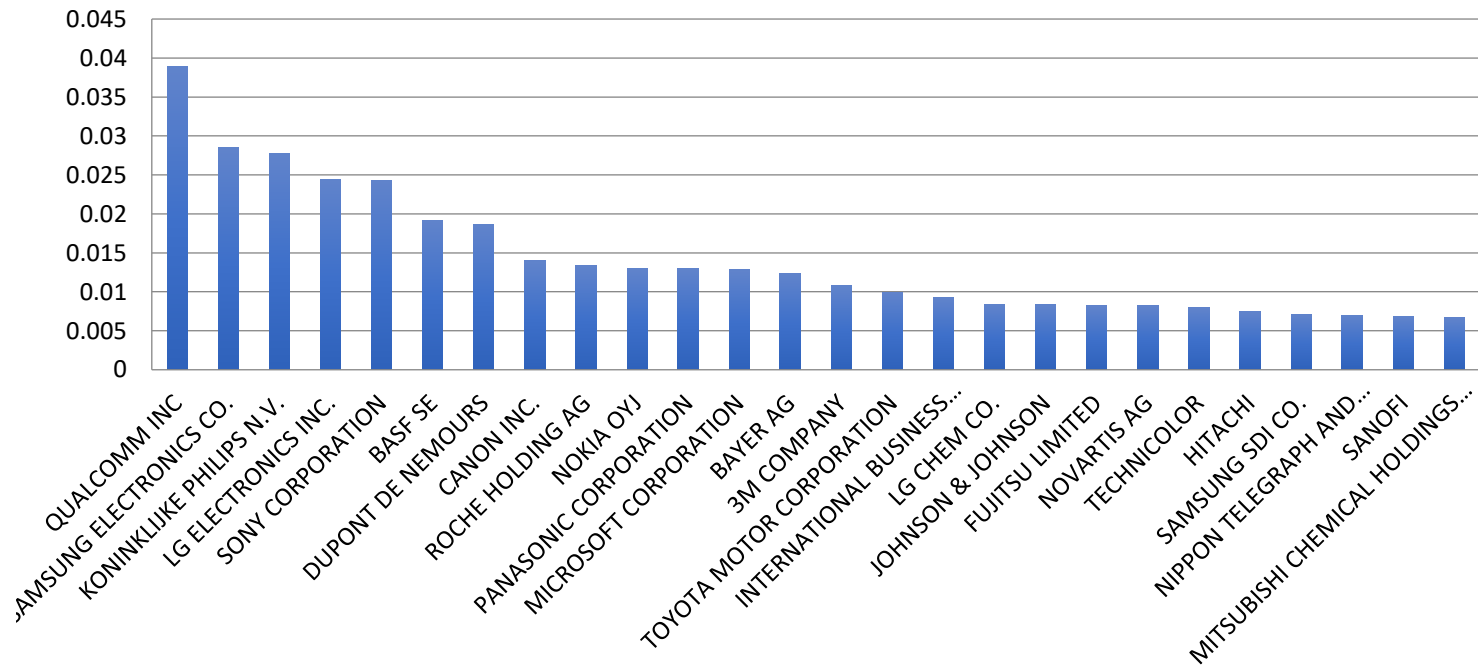
CIB2: Overview

IP5 family patents 2000-2016



CIB2: Overview

IP5 2000-2016



Next to come

- CINNOB database : the Corporate Innovation Board (firm level data) combining financial & economical data and several output data (patents, trademarks, publications, project participations)
- VICO patent database (start-up patents)
- CHEETAH patent database (fast-growing mid-size firms' patents)