

UNLOCKING THE VALUE OF PATENT DATA

WORLD INTELLECTUAL PROPERTY ORGANIZATION (WIPO)
FIRST PATENT ANALYTICS COMMUNITY OF PRACTICE (COP) ANNUAL SYMPOSIUM

17-18 SEPTEMBER 2024
GENEVA, SWITZERLAND

Frank Tietze
Professor of Innovation Engineering
frank.tietze@eng.cam.ac.uk

Head, Innovation and IP Management (IIPM) Laboratory
Centre for Technology Management (CTM)
Department of Engineering



WWW.IIPM.ENG.CAM.AC.UK



Study at CambridgeAbout the UniversityResearch at Cambridge

Quick linksSearch

Innovation and Intellectual Property Management (IIPM) Laboratory

HomeNewsResearchEducationIndustry EngagementPublicationsTeamWork with UsContactInnovation Engineering



WIPO Magazine article

2 of 5

Welcome to the Innovation and Intellectual Property (IIPM) Laboratory led by [Prof. Frank Tietze](#). We are a group of academics, students and practitioners with a deep interest in the role of intellectual property (IP) in innovation processes and systems. We are part of the University's Engineering Department, Division E (commonly known as Institute for Manufacturing - IfM), particularly the IfM's [Centre for Technology Management \(CTM\)](#).

The IIPM lab adopts an engineering management (firm level) and relational perspective on IP within distributed and collaborative (open) innovation processes and systems for emerging (manufacturing) technologies. Employing predominately a problem-driven research philosophy and empirical, mostly but not entirely qualitative methods, our research focuses on the role of IP and its strategic management for developing and deploying sustainable innovation (e.g. climate change mitigation and adaptation technologies), addressing global challenges, such as climate change, achieving sustainable development goals (SDG), and accelerating sustainability transitions.

Our research on Innovation and IP Management (IIPM) focuses on two priority areas:

News & Blog articles

Welcome to the team: Dr Soujanya Mantravadi

4 September 2024

Dr Soujanya Mantravadi joined the IIPM Lab in September to work on the ILI project (Impact Licensing Initiative) funded by EU Horizon, respectively UKRI. Soujanya brings a wealth of experience to the table and joining us from the IfM's Centre for International Manufacturing (CIM). She has a PhD in industrial engineering...

IIPM lab to embark on new, EU(UKRI)-funded project "Impact Licensing Initiative"

4 September 2024

We are pleased to announce that the IIPM lab will participate in the EU(UKRI)-funded project "Impact Licensing Initiative". Technologies and data are still underused or sub-optimally allocated to solve the societal challenges of today in Europe and beyond. Moreover, the COVID pandemic and the recent war in Ukraine has...

New working paper on the use of patent data to identify disruptive innovation

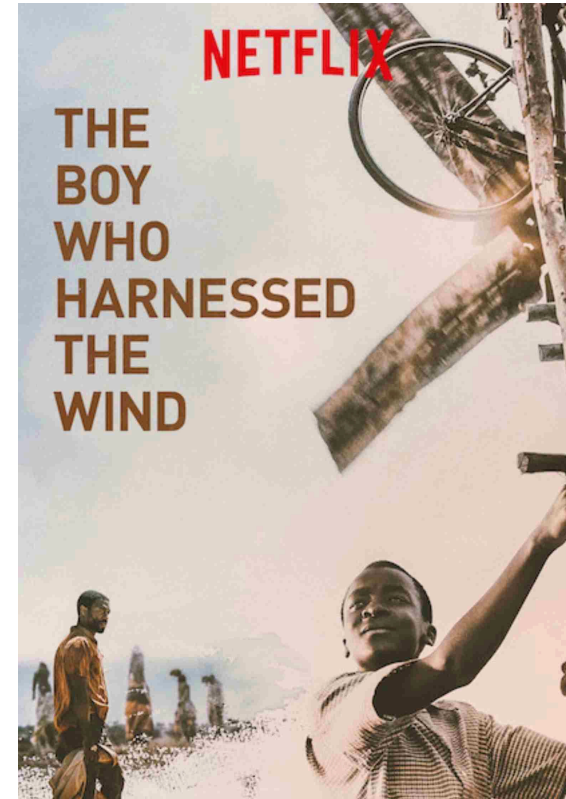


CHIEF ECONOMISTS PANEL AT RECENT EUROPEAN POLICY FOR INTELLECTUAL PROPERTY (EPIP) CONFERENCE



© F. Tietze (2024)

REMINDER: THE ORIGINAL PURPOSE OF PATENT SYSTEMS



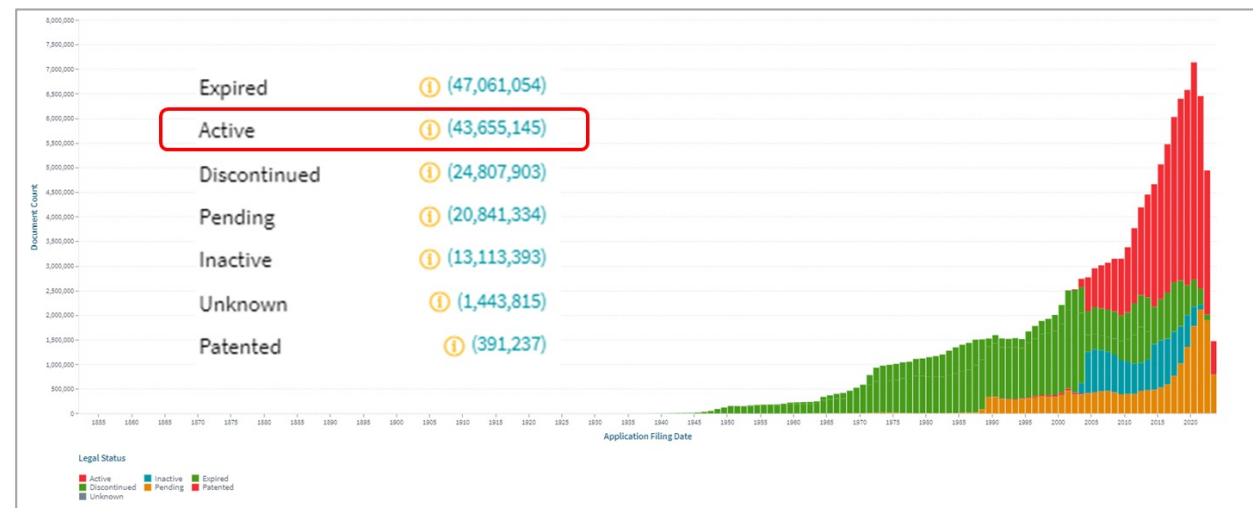
Imagine what William could have done with access to patent data...

PATENT DATABASES ARE THE OLDEST AND WORLD'S LARGEST OPEN DATA REPOSITORIES FOR TECHNICAL SOLUTION KNOWLEDGE



Image generated with <https://chatgpt.com>

85,872,443 patent families on file ¹

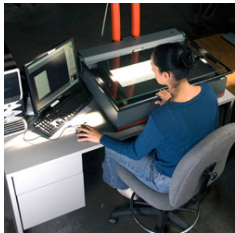


PATENT DATA PROVIDED BY THE OFFICES AND DIGITIZATION CREATED AN INDUSTRY

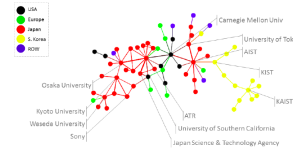
Global patent analytics market size valued at \$1.00bn (2023), projected to grow to \$3.02bn (2032) ¹



- High search costs
- Focus on individual patents
- Patent data as tool for large firms
- Large value digitising raw data
- 1984: BACON project by the trilateral offices (US, EPO, JP)



- Patent documents are mainly digital PDFs, not cloud-based
- Decreasing value in data-driven business models
- Predominance of established patent data software providers
- Limited large-scale patent data analysis capabilities



- Patent data accessible in a cloud
- Surge in innovative new entrants
- Lower search costs for accessing patent data
- Shift in value from data possession to analytical capabilities



- LLM based patent generation and analysis
- Technology race between filers and offices
- Decentralization of the IP systems allows ownership transfer tracking



<1990

~2000s

~2010s

~today

t



- Patent data fully democratized
- focus on large-scale analysis
- Significant reduction in patent data search costs via free tools
- Patent analytics accessible to non-IP experts
- New AI-compatible architectures by entrants enhance data analysis



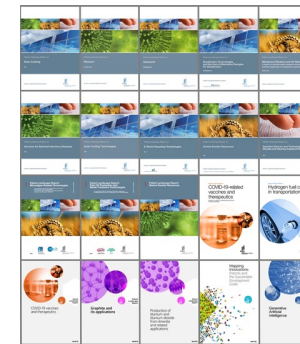
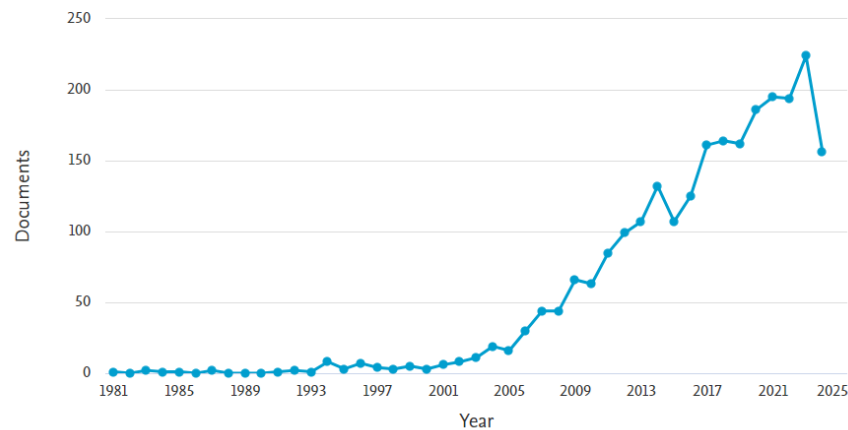
2005: My very first patent analysis ;)



DEMOCRATIZING PATENT ANALYSIS

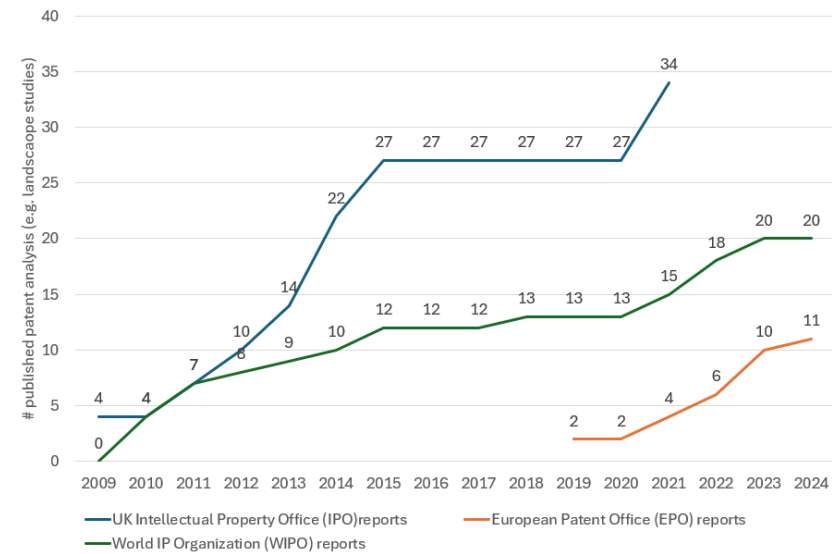
2,425 academic publications ¹

Documents by year



<https://www.epo.org/searching-for-patents/business/patent-insight-reports.html>

Patent office patent studies



HOW TO MAXIMISE THE SOCIAL VALUE OF PATENT DATA?



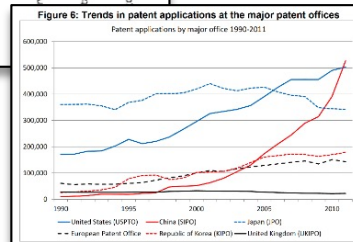
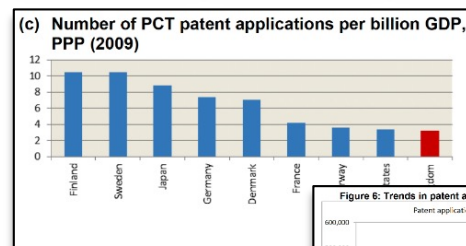
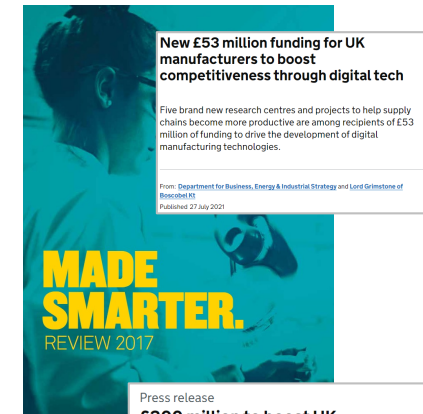
Image generated with <https://chatgpt.com>

INDIRECT USE OF PATENT ANALYSIS FOR POLICY MAKING

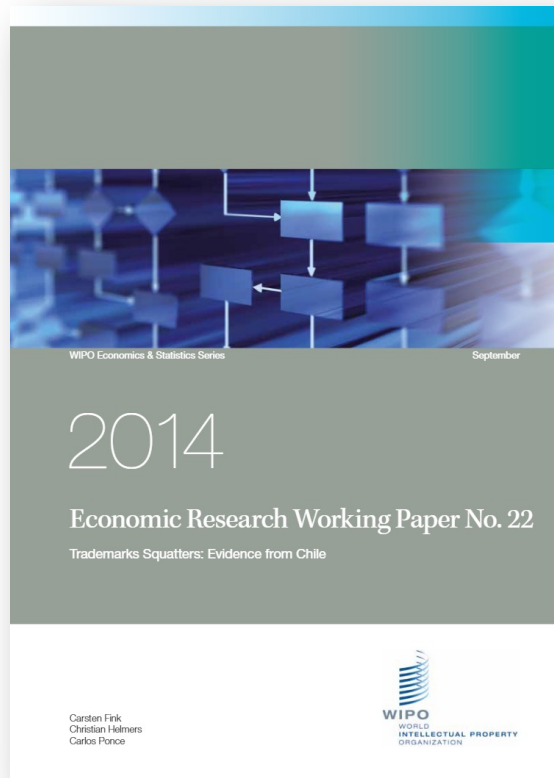
Out of 37 background studies, 6 used patent data



UK industrial and innovation policy



IMPACT OF A TRADEMARK ANALYSIS ON IP LAW



See also: Fink, C., Helmers, C., Ponce, C.J., 2018. Trademark squatters: Theory and evidence from Chile. International Journal of Industrial Organization 59, 340–371. <https://doi.org/10.1016/j.ijindorg.2018.04.004>



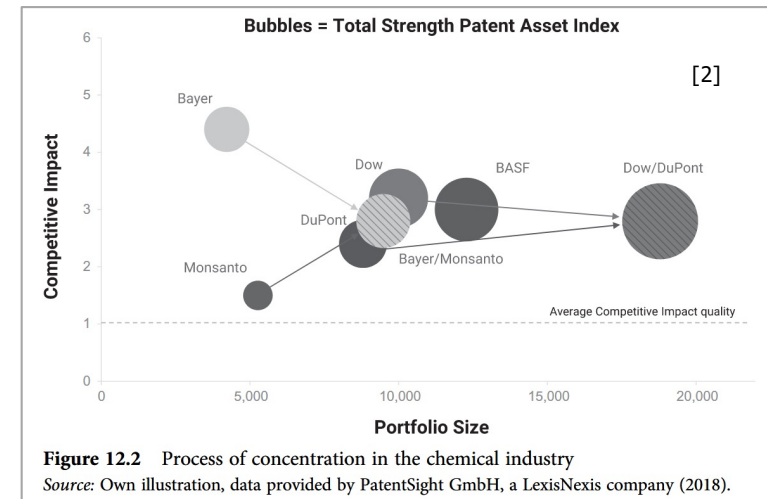
https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/new-ip-law-chile-ii-main-changes-trade-mark-system-2022-10-07_en

THE USE OF PATENT ANALYSIS IN THE DUE DILIGENCE OF MERGERS

Dow and DuPont (2017)

EC carried out patent analysis which confirmed:

1. High importance of both merging parties, and in particular one merging party, as innovators;
 2. High degree of concentration in research for new AIs (discovery stage);
 3. Significant combined share of research for new AIs accounted by the merging parties, notably in selective herbicides and insecticides; and
 4. Closeness between the merging parties in term of innovation efforts.
- Divesting a large part of DuPont's herbicide and insecticide businesses + DuPont's global R&D organisation ¹



¹ Buehler, B., Coublucq, D., Hariton, C., Langus, G., Valletti, T., 2017. Recent Developments at DG Competition: 2016/2017. Rev Ind Organ 51, 397–422.

² Illustration for Bayer/ Monsanto merger (M.8084) by Ernst, H., Guderian, C.C., Richter, M., (2022) The Innovation Environment and Knowledge Diffusion: Improving Policy Decisions through Patent Analytics, in: Taubman, A., Watal, J. (Eds.), Trade in Knowledge. Cambridge University Press, pp. 376–402.

With thanks to R. Veugelers. See also: (2019) "Innovation/IP in Competition Policy", EPIP Conference, ETH, Zurich



LEVELS OF PATENT ANALYSIS AND EXAMPLE QUESTIONS

- **Country and regions - geographical (via NUTS codes)**
(Where is development of a certain technology happening?)
- **Industry (or group of organisations)**
(Who are the major players in an industry? Which directions are they going?)
- **Technology (IPC/CPC via concordance tables)**
(Who are key developers for a certain technology?)
- **Individual entity level (Individual organisations e.g. universities, corporates, business; inventors and inventor teams)**
(What is the portfolio, inventive activity of certain entities? Is this a valuable acquisition target?)
- **Portfolio**
(Which patents are renewed? What is a portfolio worth? Which are the valuable patents?)
- **Product**
(Which patents / what IP is used in which product?)
- **Patent / IP right level**
(What are the most relevant patents in a technology? What is the value of a patent? Who would want to license this patent?)

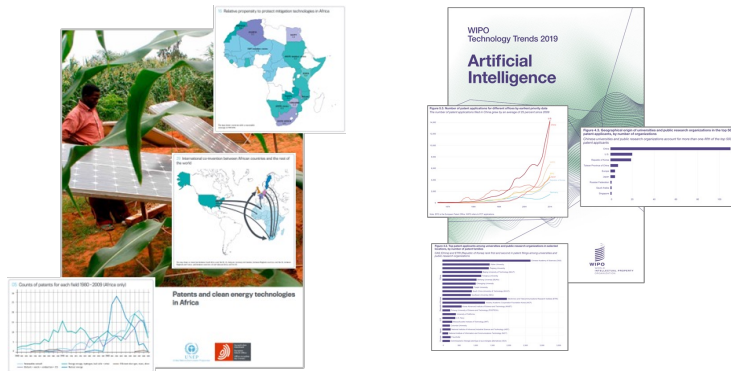


DIFFERENT TYPES OF PATENT STUDIES

- Full (dedicated) patent analytics reports vs. partial background (economic) studies (e.g. as input to a larger study)
- Descriptive patent studies (often gov reports) vs. econometric, often academic studies using patent data as one variable correlated with other variables
- “Pure” patent vs. multi-data studies (incl. patent data)
- Actor studies vs. “relational” studies (e.g. network analysis)
- Published vs. unpublished (hidden) patent analysis
- Commissioned (e.g. by EC) vs. “free” studies
- Single country vs. comparative international studies
- ...

WHERE DOES PATENT ANALYTICS START AND END? WHERE DOES THE COP WANT TO GET INVOLVED?

Landscape and trend studies
tend to present descriptive results



More advanced analysis,
e.g. knowledge flows, tech transfer, TRL estimation



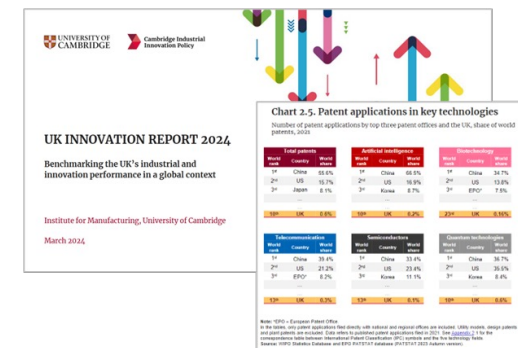
Patent data as part of
compound indicators



Studies with combined data
that require dataset matching



The many innovation policy/ foresight reports
that include some form of patent analysis



INCREASINGLY DATA IS AVAILABLE FOR OTHER IP RIGHTS THAT COULD BE ANALYSED

Trademarks



Geographical indications

	2020	2012
Italy	297	248
France	249	192
Spain	192	161
Portugal	138	118
Greece	104	97
Germany	91	89
UK	69	46
Poland	31	35
Czech Republic	29	28
Croatia	24	0
Slovenia	22	16
Austria	15	14
Belgium	15	13
Hungary	14	12
Slovakia	12	10
the Netherlands	11	9
Denmark	8	5
Finland	7	8
Ireland	7	4
Lithuania	7	2
Romania	7	1
Sweden	6	6
Cyprus	5	2
Luxembourg	4	4
Bulgaria	3	1
Latvia	3	0
Andorra	1	0

Note(s): Blocks represented by colour, South, North, Central East

Joose, S., Olders, P., Boonstra, W.J., 2021. Why are geographical indications unevenly distributed over Europe? British Food Journal 123, 490–510.

Design rights

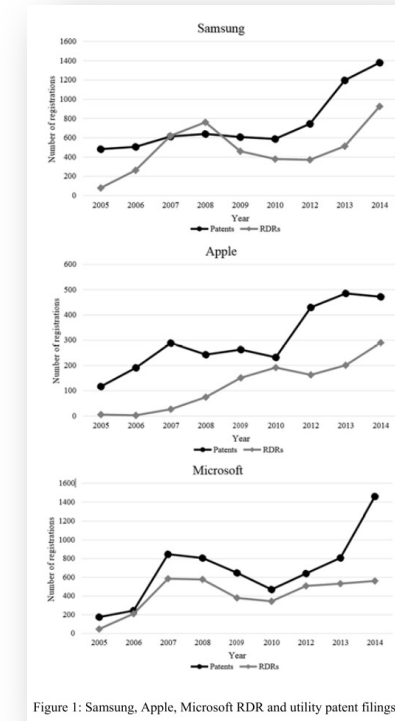


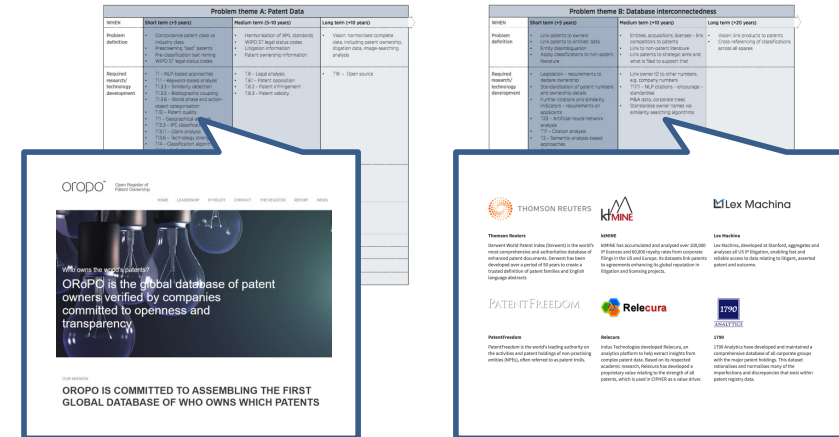
Figure 1: Samsung, Apple, Microsoft RDR and utility patent filings

Wolf, P., Tietze, F., Schweisfurth, T., Moultrie, J. (2017). Registered Design Rights as Innovation Indicators. R&D Management Conference. Leuven, Belgium.

REFLECTING ON THE FIVE PROBLEMS WE IDENTIFIED IN 2017



EXPLORING THE FUTURE OF PATENT ANALYTICS



PATENT ANALYTICS PROBLEM THEMES AND DISCUSSION

Problem A – Patent data

Problem B – Patent database interconnectedness

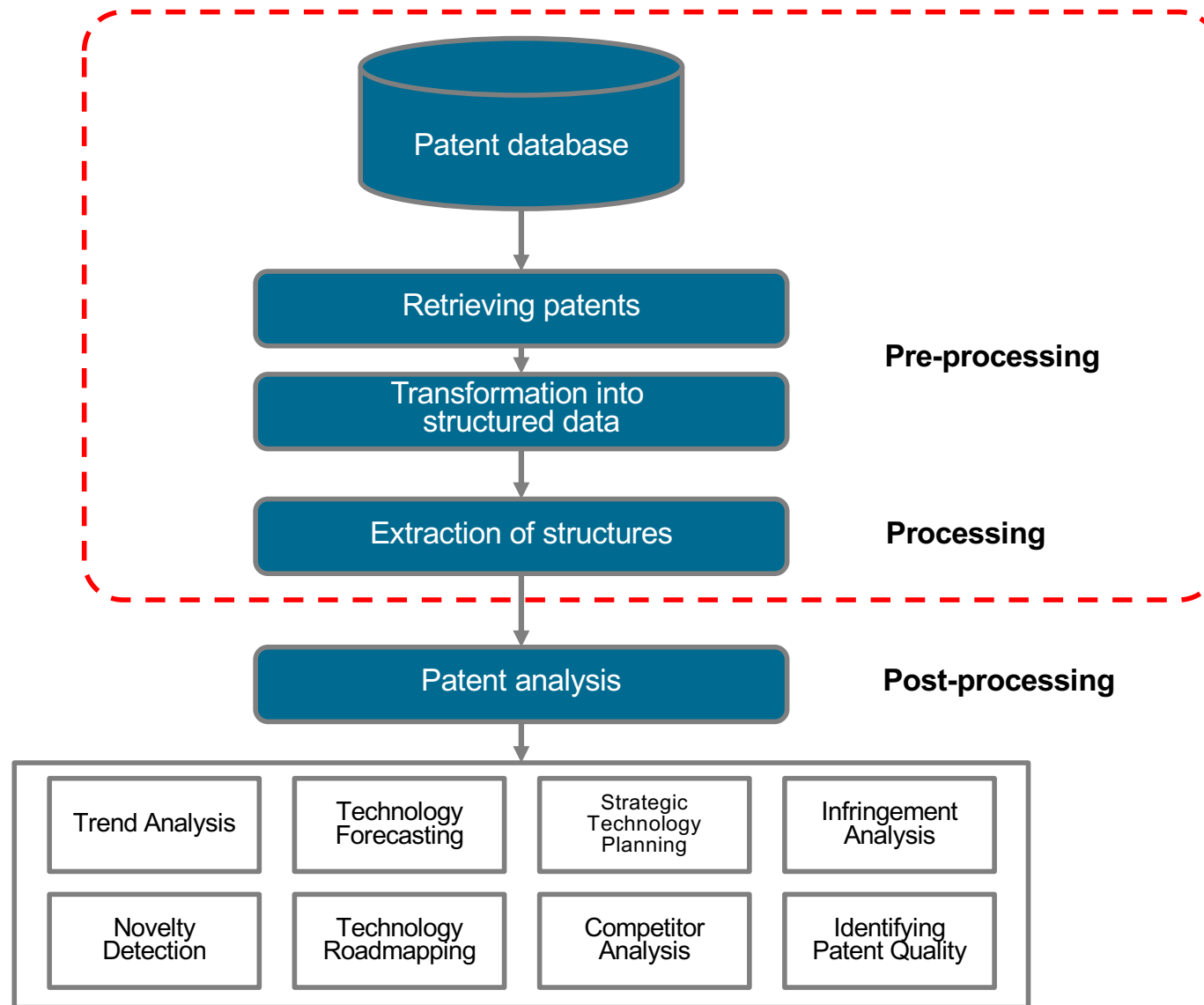
Problem C – Patent data analysis

Problem D – Patent information visualisation

Problem E – Patent quality



THE PATENT ANALYSIS WORK FLOW



DATA QUALITY AND THE NEED FOR CLEANSING – EXAMPLE APPLICANT NAME

Nowadays, databases typically employ “corporate tree” functions

Table 8.4 - Examples of consolidated assignee names

Original Assignee Name	Homogenized / Consolidated Original Assignee Name	OATY*
Li Ming Chiang, US	Ming-Chiang Li	1
Li; Ming-Chiang	Ming-Chiang Li	1
Min-Chiang Li	Ming-Chiang Li	1
Arco Chem Tech, US	Unilever	4
CPC International Inc, US		
Arco Chem Tech, US	Unilever	4
CPC International Inc.	Unilever	4
CPC International Inc, US	Unilever	4
Arco Chem Tech, US		
CPC International, Inc.	Unilever	4
Arco Chemical Technology, Inc.		
Bestfoods	Unilever	4

Note: * 1 (individual inventor); 4 (large firm)

Tietze, F. (2012). Technology Market Transactions - Auctions, Intermediaries and Innovation. Cheltenham Edgar Elgar Publishing.

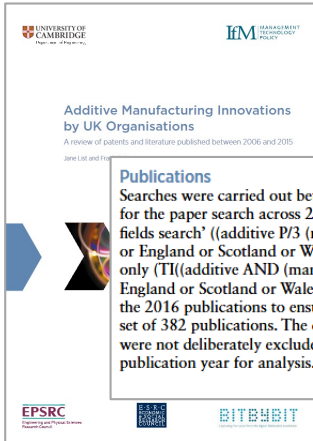
Table 12: Top 10 patenting organizations with patent count and ranking before and after harmonization.

Harmonized Name (after second round of Harmonizing)	After harmonization		Before harmonization		Improvement
	# patents	Rank	# patents	Rank	
MATSUSHITA ELECTRIC INDUSTRIAL COMPANY	442.211	1	326.425	1	35,47%
NEC CORPORATION	347.687	2	184.195	7	88,76%
HITACHI	342.476	3	260.455	2	31,49%
TOSHIBA CORPORATION	336.649	4	236.744	3	42,20%
CANON	334.891	5	202.820	4	65,12%
MITSUBISHI ELECTRIC CORPORATION	305.575	6	187.569	6	62,91%
SAMSUNG ELECTRONICS COMPANY	274.666	7	201.932	5	36,02%
FUJITSU	270.722	8	158.045	8	71,29%
SONY CORPORATION	258.811	9	144.891	9	78,62%
SIEMENS	256.874	10	104.848	15	145,00%

Peeters, B., Song, X., Callaert, J., Grouwels, J., & Van Looy, B. (2009). Harmonizing harmonized patentee names: an exploratory assessment of top patentees. Luxembourg: EUROSTAT working paper and Studies.

BUT: We still do not know who owns which patents (and which are traded/ licensed), but only who registered them in the first place!

DEVELOPING EFFECTIVE SEARCH STRATEGIES AND COMPILING RELIABLE DATASETS REMAINS COSTLY



Publications

Searches were carried out between 27 March and 24 April 2017. A high-precision search string was used for the paper search across 23 databases, consistent with the approach taken for the patent search. An 'all fields search' ((additive P/3 (manufact* or fabricat*)) OR (3D P/3 PRINT*)) AND af(UK or GB or Britain or England or Scotland or Wales) resulted in 3416 documents, which was then reduced to a 'title' search only (TI((additive AND (manufact* or fabricat*)) or (3D AND print*)) AND AF(UK or GB or Britain or England or Scotland or Wales)). This resulted in 589 publications after removal of duplicates. We excluded the 2016 publications to ensure we covered the same period as for patent documents. This resulted in a set of 382 publications. The dataset includes dissertations, which were not specifically searched for, but were not deliberately excluded if the search picked them up. Papers were downloaded from ProQuest by publication year for analysis.



Fuel cells in transport

Part 1

(Tag: "SOFC Fuel Cells", "DMFC", "DAFC", "Direct or Reforming Fuel Cells", "AMFC Alkaline Membrane Fuel Cells", "PMFC Phosphoric Acid Fuel Cells", "MCFC Molten Carbonate Fuel Cells", "Fuel Cells", "REM Fuel Cells", "Fuel Cell Manufacturing, Stacking" AND CPC=(Y02O 90/00, Y02O 90/04, Y02O 90/10, Y02O 90/14, Y02O 90/16, Y02O 90/18, Y02O 90/20, Y02O 90/22, Y02O 90/24, Y02O 90/26, Y02O 90/28, Y02O 90/30, Y02O 90/32, Y02O 90/34, Y02O 90/36, Y02O 90/38, Y02O 90/40, Y02O 90/42, Y02O 90/44, Y02O 90/46, Y02O 90/48, Y02O 90/50, Y02O 90/52, Y02O 90/54, Y02O 90/56, Y02O 90/58, Y02O 90/60, Y02O 90/62, Y02O 90/64, Y02O 90/66, Y02O 90/68, Y02O 90/70, Y02O 90/72, Y02O 90/74, Y02O 90/76, Y02O 90/78, Y02O 90/80, Y02O 90/82, Y02O 90/84, Y02O 90/86, Y02O 90/88, Y02O 90/90, Y02O 90/92, Y02O 90/94, Y02O 90/96, Y02O 90/98, Y02O 90/100, Y02O 90/102, Y02O 90/104, Y02O 90/106, Y02O 90/108, Y02O 90/110, Y02O 90/112, Y02O 90/114, Y02O 90/116, Y02O 90/118, Y02O 90/120, Y02O 90/122, Y02O 90/124, Y02O 90/126, Y02O 90/128, Y02O 90/130, Y02O 90/132, Y02O 90/134, Y02O 90/136, Y02O 90/138, Y02O 90/140, Y02O 90/142, Y02O 90/144, Y02O 90/146, Y02O 90/148, Y02O 90/150, Y02O 90/152, Y02O 90/154, Y02O 90/156, Y02O 90/158, Y02O 90/160, Y02O 90/162, Y02O 90/164, Y02O 90/166, Y02O 90/168, Y02O 90/170, Y02O 90/172, Y02O 90/174, Y02O 90/176, Y02O 90/178, Y02O 90/180, Y02O 90/182, Y02O 90/184, Y02O 90/186, Y02O 90/188, Y02O 90/190, Y02O 90/192, Y02O 90/194, Y02O 90/196, Y02O 90/198, Y02O 90/200, Y02O 90/202, Y02O 90/204, Y02O 90/206, Y02O 90/208, Y02O 90/210, Y02O 90/212, Y02O 90/214, Y02O 90/216, Y02O 90/218, Y02O 90/220, Y02O 90/222, Y02O 90/224, Y02O 90/226, Y02O 90/228, Y02O 90/230, Y02O 90/232, Y02O 90/234, Y02O 90/236, Y02O 90/238, Y02O 90/240, Y02O 90/242, Y02O 90/244, Y02O 90/246, Y02O 90/248, Y02O 90/250, Y02O 90/252, Y02O 90/254, Y02O 90/256, Y02O 90/258, Y02O 90/260, Y02O 90/262, Y02O 90/264, Y02O 90/266, Y02O 90/268, Y02O 90/270, Y02O 90/272, Y02O 90/274, Y02O 90/276, Y02O 90/278, Y02O 90/280, Y02O 90/282, Y02O 90/284, Y02O 90/286, Y02O 90/288, Y02O 90/290, Y02O 90/292, Y02O 90/294, Y02O 90/296, Y02O 90/298, Y02O 90/300, Y02O 90/302, Y02O 90/304, Y02O 90/306, Y02O 90/308, Y02O 90/310, Y02O 90/312, Y02O 90/314, Y02O 90/316, Y02O 90/318, Y02O 90/320, Y02O 90/322, Y02O 90/324, Y02O 90/326, Y02O 90/328, Y02O 90/330, Y02O 90/332, Y02O 90/334, Y02O 90/336, Y02O 90/338, Y02O 90/340, Y02O 90/342, Y02O 90/344, Y02O 90/346, Y02O 90/348, Y02O 90/350, Y02O 90/352, Y02O 90/354, Y02O 90/356, Y02O 90/358, Y02O 90/360, Y02O 90/362, Y02O 90/364, Y02O 90/366, Y02O 90/368, Y02O 90/370, Y02O 90/372, Y02O 90/374, Y02O 90/376, Y02O 90/378, Y02O 90/380, Y02O 90/382, Y02O 90/384, Y02O 90/386, Y02O 90/388, Y02O 90/390, Y02O 90/392, Y02O 90/394, Y02O 90/396, Y02O 90/398, Y02O 90/400, Y02O 90/402, Y02O 90/404, Y02O 90/406, Y02O 90/408, Y02O 90/410, Y02O 90/412, Y02O 90/414, Y02O 90/416, Y02O 90/418, Y02O 90/420, Y02O 90/422, Y02O 90/424, Y02O 90/426, Y02O 90/428, Y02O 90/430, Y02O 90/432, Y02O 90/434, Y02O 90/436, Y02O 90/438, Y02O 90/440, Y02O 90/442, Y02O 90/444, Y02O 90/446, Y02O 90/448, Y02O 90/450, Y02O 90/452, Y02O 90/454, Y02O 90/456, Y02O 90/458, Y02O 90/460, Y02O 90/462, Y02O 90/464, Y02O 90/466, Y02O 90/468, Y02O 90/470, Y02O 90/472, Y02O 90/474, Y02O 90/476, Y02O 90/478, Y02O 90/480, Y02O 90/482, Y02O 90/484, Y02O 90/486, Y02O 90/488, Y02O 90/490, Y02O 90/492, Y02O 90/494, Y02O 90/496, Y02O 90/498, Y02O 90/500, Y02O 90/502, Y02O 90/504, Y02O 90/506, Y02O 90/508, Y02O 90/510, Y02O 90/512, Y02O 90/514, Y02O 90/516, Y02O 90/518, Y02O 90/520, Y02O 90/522, Y02O 90/524, Y02O 90/526, Y02O 90/528, Y02O 90/530, Y02O 90/532, Y02O 90/534, Y02O 90/536, Y02O 90/538, Y02O 90/540, Y02O 90/542, Y02O 90/544, Y02O 90/546, Y02O 90/548, Y02O 90/550, Y02O 90/552, Y02O 90/554, Y02O 90/556, Y02O 90/558, Y02O 90/560, Y02O 90/562, Y02O 90/564, Y02O 90/566, Y02O 90/568, Y02O 90/570, Y02O 90/572, Y02O 90/574, Y02O 90/576, Y02O 90/578, Y02O 90/580, Y02O 90/582, Y02O 90/584, Y02O 90/586, Y02O 90/588, Y02O 90/590, Y02O 90/592, Y02O 90/594, Y02O 90/596, Y02O 90/598, Y02O 90/600, Y02O 90/602, Y02O 90/604, Y02O 90/606, Y02O 90/608, Y02O 90/610, Y02O 90/612, Y02O 90/614, Y02O 90/616, Y02O 90/618, Y02O 90/620, Y02O 90/622, Y02O 90/624, Y02O 90/626, Y02O 90/628, Y02O 90/630, Y02O 90/632, Y02O 90/634, Y02O 90/636, Y02O 90/638, Y02O 90/640, Y02O 90/642, Y02O 90/644, Y02O 90/646, Y02O 90/648, Y02O 90/650, Y02O 90/652, Y02O 90/654, Y02O 90/656, Y02O 90/658, Y02O 90/660, Y02O 90/662, Y02O 90/664, Y02O 90/666, Y02O 90/668, Y02O 90/670, Y02O 90/672, Y02O 90/674, Y02O 90/676, Y02O 90/678, Y02O 90/680, Y02O 90/682, Y02O 90/684, Y02O 90/686, Y02O 90/688, Y02O 90/690, Y02O 90/692, Y02O 90/694, Y02O 90/696, Y02O 90/698, Y02O 90/700, Y02O 90/702, Y02O 90/704, Y02O 90/706, Y02O 90/708, Y02O 90/710, Y02O 90/712, Y02O 90/714, Y02O 90/716, Y02O 90/718, Y02O 90/720, Y02O 90/722, Y02O 90/724, Y02O 90/726, Y02O 90/728, Y02O 90/730, Y02O 90/732, Y02O 90/734, Y02O 90/736, Y02O 90/738, Y02O 90/740, Y02O 90/742, Y02O 90/744, Y02O 90/746, Y02O 90/748, Y02O 90/750, Y02O 90/752, Y02O 90/754, Y02O 90/756, Y02O 90/758, Y02O 90/760, Y02O 90/762, Y02O 90/764, Y02O 90/766, Y02O 90/768, Y02O 90/770, Y02O 90/772, Y02O 90/774, Y02O 90/776, Y02O 90/778, Y02O 90/780, Y02O 90/782, Y02O 90/784, Y02O 90/786, Y02O 90/788, Y02O 90/790, Y02O 90/792, Y02O 90/794, Y02O 90/796, Y02O 90/798, Y02O 90/800, Y02O 90/802, Y02O 90/804, Y02O 90/806, Y02O 90/808, Y02O 90/810, Y02O 90/812, Y02O 90/814, Y02O 90/816, Y02O 90/818, Y02O 90/820, Y02O 90/822, Y02O 90/824, Y02O 90/826, Y02O 90/828, Y02O 90/830, Y02O 90/832, Y02O 90/834, Y02O 90/836, Y02O 90/838, Y02O 90/840, Y02O 90/842, Y02O 90/844, Y02O 90/846, Y02O 90/848, Y02O 90/850, Y02O 90/852, Y02O 90/854, Y02O 90/856, Y02O 90/858, Y02O 90/860, Y02O 90/862, Y02O 90/864, Y02O 90/866, Y02O 90/868, Y02O 90/870, Y02O 90/872, Y02O 90/874, Y02O 90/876, Y02O 90/878, Y02O 90/880, Y02O 90/882, Y02O 90/884, Y02O 90/886, Y02O 90/888, Y02O 90/890, Y02O 90/892, Y02O 90/894, Y02O 90/896, Y02O 90/898, Y02O 90/900, Y02O 90/902, Y02O 90/904, Y02O 90/906, Y02O 90/908, Y02O 90/910, Y02O 90/912, Y02O 90/914, Y02O 90/916, Y02O 90/918, Y02O 90/920, Y02O 90/922, Y02O 90/924, Y02O 90/926, Y02O 90/928, Y02O 90/930, Y02O 90/932, Y02O 90/934, Y02O 90/936, Y02O 90/938, Y02O 90/940, Y02O 90/942, Y02O 90/944, Y02O 90/946, Y02O 90/948, Y02O 90/950, Y02O 90/952, Y02O 90/954, Y02O 90/956, Y02O 90/958, Y02O 90/960, Y02O 90/962, Y02O 90/964, Y02O 90/966, Y02O 90/968, Y02O 90/970, Y02O 90/972, Y02O 90/974, Y02O 90/976, Y02O 90/978, Y02O 90/980, Y02O 90/982, Y02O 90/984, Y02O 90/986, Y02O 90/988, Y02O 90/990, Y02O 90/992, Y02O 90/994, Y02O 90/996, Y02O 90/998, Y02O 90/1000, Y02O 90/1002, Y02O 90/1004, Y02O 90/1006, Y02O 90/1008, Y02O 90/1010, Y02O 90/1012, Y02O 90/1014, Y02O 90/1016, Y02O 90/1018, Y02O 90/1020, Y02O 90/1022, Y02O 90/1024, Y02O 90/1026, Y02O 90/1028, Y02O 90/1030, Y02O 90/1032, Y02O 90/1034, Y02O 90/1036, Y02O 90/1038, Y02O 90/1040, Y02O 90/1042, Y02O 90/1044, Y02O 90/1046, Y02O 90/1048, Y02O 90/1050, Y02O 90/1052, Y02O 90/1054, Y02O 90/1056, Y02O 90/1058, Y02O 90/1060, Y02O 90/1062, Y02O 90/1064, Y02O 90/1066, Y02O 90/1068, Y02O 90/1070, Y02O 90/1072, Y02O 90/1074, Y02O 90/1076, Y02O 90/1078, Y02O 90/1080, Y02O 90/1082, Y02O 90/1084, Y02O 90/1086, Y02O 90/1088, Y02O 90/1090, Y02O 90/1092, Y02O 90/1094, Y02O 90/1096, Y02O 90/1098, Y02O 90/1100, Y02O 90/1102, Y02O 90/1104, Y02O 90/1106, Y02O 90/1108, Y02O 90/1110, Y02O 90/1112, Y02O 90/1114, Y02O 90/1116, Y02O 90/1118, Y02O 90/1120, Y02O 90/1122, Y02O 90/1124, Y02O 90/1126, Y02O 90/1128, Y02O 90/1130, Y02O 90/1132, Y02O 90/1134, Y02O 90/1136, Y02O 90/1138, Y02O 90/1140, Y02O 90/1142, Y02O 90/1144, Y02O 90/1146, Y02O 90/1148, Y02O 90/1150, Y02O 90/1152, Y02O 90/1154, Y02O 90/1156, Y02O 90/1158, Y02O 90/1160, Y02O 90/1162, Y02O 90/1164, Y02O 90/1166, Y02O 90/1168, Y02O 90/1170, Y02O 90/1172, Y02O 90/1174, Y02O 90/1176, Y02O 90/1178, Y02O 90/1180, Y02O 90/1182, Y02O 90/1184, Y02O 90/1186, Y02O 90/1188, Y02O 90/1190, Y02O 90/1192, Y02O 90/1194, Y02O 90/1196, Y02O 90/1198, Y02O 90/1200, Y02O 90/1202, Y02O 90/1204, Y02O 90/1206, Y02O 90/1208, Y02O 90/1210, Y02O 90/1212, Y02O 90/1214, Y02O 90/1216, Y02O 90/1218, Y02O 90/1220, Y02O 90/1222, Y02O 90/1224, Y02O 90/1226, Y02O 90/1228, Y02O 90/1230, Y02O 90/1232, Y02O 90/1234, Y02O 90/1236, Y02O 90/1238, Y02O 90/1240, Y02O 90/1242, Y02O 90/1244, Y02O 90/1246, Y02O 90/1248, Y02O 90/1250, Y02O 90/1252, Y02O 90/1254, Y02O 90/1256, Y02O 90/1258, Y02O 90/1260, Y02O 90/1262, Y02O 90/1264, Y02O 90/1266, Y02O 90/1268, Y02O 90/1270, Y02O 90/1272, Y02O 90/1274, Y02O 90/1276, Y02O 90/1278, Y02O 90/1280, Y02O 90/1282, Y02O 90/1284, Y02O 90/1286, Y02O 90/1288, Y02O 90/1290, Y02O 90/1292, Y02O 90/1294, Y02O 90/1296, Y02O 90/1298, Y02O 90/1300, Y02O 90/1302, Y02O 90/1304, Y02O 90/1306, Y02O 90/1308, Y02O 90/1310, Y02O 90/1312, Y02O 90/1314, Y02O 90/1316, Y02O 90/1318, Y02O 90/1320, Y02O 90/1322, Y02O 90/1324, Y02O 90/1326, Y02O 90/1328, Y02O 90/1330, Y02O 90/1332, Y02O 90/1334, Y02O 90/1336, Y02O 90/1338, Y02O 90/1340, Y02O 90/1342, Y02O 90/1344, Y02O 90/1346, Y02O 90/1348, Y02O 90/1350, Y02O 90/1352, Y02O 90/1354, Y02O 90/1356, Y02O 90/1358, Y02O 90/1360, Y02O 90/1362, Y02O 90/1364, Y02O 90/1366, Y02O 90/1368, Y02O 90/1370, Y02O 90/1372, Y02O 90/1374, Y02O 90/1376, Y02O 90/1378, Y02O 90/1380, Y02O 90/1382, Y02O 90/1384, Y02O 90/1386, Y02O 90/1388, Y02O 90/1390, Y02O 90/1392, Y02O 90/1394, Y02O 90/1396, Y02O 90/1398, Y02O 90/1400, Y02O 90/1402, Y02O 90/1404, Y02O 90/1406, Y02O 90/1408, Y02O 90/1410, Y02O 90/1412, Y02O 90/1414, Y02O 90/1416, Y02O 90/1418, Y02O 90/1420, Y02O 90/1422, Y02O 90/1424, Y02O 90/1426, Y02O 90/1428, Y02O 90/1430, Y02O 90/1432, Y02O 90/1434, Y02O 90/1436, Y02O 90/1438, Y02O 90/1440, Y02O 90/1442, Y02O 90/1444, Y02O 90/1446, Y02O 90/1448, Y02O 90/1450, Y02O 90/1452, Y02O 90/1454, Y02O 90/1456, Y02O 90/1458, Y02O 90/1460, Y02O 90/1462, Y02O 90/1464, Y02O 90/1466, Y02O 90/1468, Y02O 90/1470, Y02O 90/1472, Y02O 90/1474, Y02O 90/1476, Y02O 90/1478, Y02O 90/1480, Y02O 90/1482, Y02O 90/1484, Y02O 90/1486, Y02O 90/1488, Y02O 90/1490, Y02O 90/1492, Y02O 90/1494, Y02O 90/1496, Y02O 90/1498, Y02O 90/1500, Y02O 90/1502, Y02O 90/1504, Y02O 90/1506, Y02O 90/1508, Y02O 90/1510, Y02O 90/1512, Y02O 90/1514, Y02O 90/1516, Y02O 90/1518, Y02O 90/1520, Y02O 90/1522, Y02O 90/1524, Y02O 90/1526, Y02O 90/1528, Y02O 90/1530, Y02O 90/1532, Y02O 90/1534, Y02O 90/1536, Y02O 90/1538, Y02O 90/1540, Y02O 90/1542, Y02O 90/1544, Y02O 90/1546, Y02O 90/1548, Y02O 90/1550, Y02O 90/1552, Y02O 90/1554, Y02O 90/1556, Y02O 90/1558, Y02O 90/1560, Y02O 90/1562, Y02O 90/1564, Y02O 90/1566, Y02O 90/1568, Y02O 90/1570, Y02O 90/1572, Y02O 90/1574, Y02O 90/1576, Y02O 90/1578, Y02O 90/1580, Y02O 90/1582, Y02O 90/1584, Y02O 90/1586, Y02O 90/1588, Y02O 90/1590, Y02O 90/1592, Y02O 90/1594, Y02O 90/1596, Y02O 90/1598, Y02O 90/1600, Y02O 90/1602, Y02O 90/1604, Y02O 90/1606, Y02O 90/1608, Y02O 90/1610, Y02O 90/1612, Y02O 90/1614, Y02O 90/1616, Y02O 90/1618, Y02O 90/1620, Y02O 90/1622, Y02O 90/1624, Y02O 90/1626, Y02O 90/1628, Y02O 90/1630, Y02O 90/1632, Y02O 90/1634, Y02O 90/1636, Y02O 90/1638, Y02O 90/1640, Y02O 90/1642, Y02O 90/1644, Y02O 90/1646, Y02O 90/1648, Y02O 90/1650, Y02O 90/1652, Y02O 90/1654, Y02O 90/1656, Y02O 90/1658, Y02O 90/1660, Y02O 90/1662, Y02O 90/1664, Y02O 90/1666, Y02O 90/1668, Y02O 90/1670, Y02O 90/1672, Y02O 90/1674, Y02O 90/1676, Y02O 90/1678, Y02O 90/1680, Y02O 90/1682, Y02O 90/1684, Y02O 90/1686, Y02O 90/1688, Y02O 90/1690, Y02O 90/1692, Y02O 90/1694, Y02O 90/1696, Y02O 90/1698, Y02O 90/1700, Y02O 90/1702, Y02O 90/1704, Y02O 90/1706, Y02O 90/1708, Y02O 90/1710, Y02O 90/1712, Y02O 90/1714, Y02O 90/1716, Y02O 90/1718, Y02O 90/1720, Y02O 90/1722, Y02O 90/1724, Y02O 90/1726, Y02O 90/1728, Y02O 90/1730, Y02O 90/1732, Y02O 90/1734, Y02O 90/1736, Y02O 90/1738, Y02O 90/1740, Y02O 90/1742, Y02O 90/1744, Y02O 90/1746, Y02O 90/1748, Y02O 90/1750, Y02O 90/1752, Y02O 90/1754, Y02O 90/1756, Y02O 90/1758, Y02O 90/1760, Y02O 90/1762, Y02O 90/1764, Y02O 90/1766, Y02O 90/1768, Y02O 90/1770, Y02O 90/1772, Y02O 90/1774, Y02O 90/1776, Y02O 90/1778, Y02O 90/1780, Y02O 90/1782, Y02O 90/1784, Y02O 90/1786, Y02O 90/1788, Y02O 90/1790, Y02O 90/1792, Y02O 90/1794, Y02O 90/1796, Y02O 90/1798, Y02O 90/1800, Y02O 90/1802, Y02O 90/1804, Y02O 90/1806, Y02O 90/1808, Y02O 90/1810, Y02O 90/1812, Y02O 90/1814, Y02O 90/1816, Y02O 90/1818, Y02O 90/1820, Y02O 90/1822, Y02O 90/1824, Y02O 90/1826, Y02O 90/1828, Y02O 90/1830, Y02O 90/1832, Y02O 90/1834, Y02O 90/1836, Y02O 90/1838, Y02O 90/1840, Y02O 90/1842, Y02O 90/1844, Y02O 90/1846, Y02O 90/1848, Y02O 90/1850, Y02O 90/1852, Y02O 90/1854, Y02O 90/1856, Y02O 90/1858, Y02O 90/1860, Y02O 90/1862, Y02O 90/1864, Y02O 90/1866, Y02O 90/1868, Y02O 90/1870, Y02O 90/1872, Y02O 90/1874, Y02O 90/1876, Y02O 90/1878, Y02O 90/1880, Y02O 90/1882, Y02O 90/1884, Y02O 90/1886, Y02O 90/1888, Y02O 90



INCOMPLETENESS OF DATA REMAINS A PROBLEM

- Y02 data based study of climate adaptation technology contributions
- Patstat ... based on inventor residency data we aggregated the data into four groups, i.e. high-income, upper-middle income, lower-middle income, and low-income countries ... World Bank country classification
- “inventor residence is not always reported for all patents, resulting in a bias towards countries and patent offices that report inventor residence data more accurately.”

Table 2: Inventor country patenting activity in adaptation technologies between 1980 and 2019

No	Global	Country	Region	Coastal (Y02A10) [%]	Water (Y02A20) [%]	Infrastructure (Y02A30) [%]	Agriculture (Y02A40) [%]	Health (Y02A50) [%]	Indirect (Y02A60) [%]	Total Patents	Patents / population ¹	Patents / GDP ²
High-income countries												
High-income overall				<u>2.45</u>	<u>14.53</u>	<u>8.32</u>	<u>24.15</u>	<u>44.59</u>	<u>5.93</u>	<u>57.781</u>	<u>5.22</u>	<u>0.90</u>
1	1	United States	North America	0.75	6.93	5.31	17.28	58.06	11.68	19,310	5.83	0.92
2	2	Germany	Europe & Central Asia	1.19	9.54	13.11	16.34	56.22	3.65	6,468	7.78	1.66
3	3	Japan	East Asia & Pacific	1.40	10.18	15.25	18.05	49.58	5.51	6,444	5.10	1.28
4	4	France	Europe & Central Asia	1.28	6.41	9.58	20.71	55.39	6.64	3,356	4.97	1.27
5	6	United Kingdom	Europe & Central Asia	1.13	9.32	6.84	16.28	52.16	7.29	3,100	4.62	1.15
6	7	Korea, Rep.	East Asia & Pacific	3.85	11.68	11.00	24.49	45.01	3.96	2,728	5.28	1.66
7	8	Canada	North America	1.48	10.20	5.10	25.65	49.62	7.96	1,961	5.16	1.19
8	9	Netherlands	Europe & Central Asia	1.84	8.01	8.50	44.93	32.74	3.97	1,411	8.09	1.55
9	10	Switzerland	Europe & Central Asia	1.73	6.68	10.58	17.70	59.04	4.28	1,333	15.43	1.80
10	11	Italy	Europe & Central Asia	2.34	11.19	11.41	21.16	51.17	2.72	1,323	2.23	0.70
Remaining high-income countries (68 in total)				2.62	15.48	8.08	24.47	43.39	5.96	10,347	5.00	0.82
Upper-middle income countries												
Upper-middle income overall				<u>1.71</u>	<u>15.11</u>	<u>4.67</u>	<u>26.10</u>	<u>49.63</u>	<u>2.78</u>	<u>4,926</u>	<u>0.17</u>	<u>0.28</u>
1	5	China	East Asia & Pacific	2.35	12.32	11.74	26.17	42.51	4.92	3,110	0.22	0.21
2	23	Brazil	Latin America & Caribbean	0.79	12.37	2.89	28.62	51.32	4.21	389	0.18	0.26
3	25	Russian Federation	Europe & Central Asia	1.20	21.26	8.38	22.46	38.62	8.08	334	0.23	0.22
4	28	Mexico	Latin America & Caribbean	2.86	15.43	10.29	24.00	44.57	2.88	175	0.14	0.16
5	29	South Africa	Sub-Saharan Africa	3.43	18.06	6.66	25.14	42.29	3.43	175	0.30	0.52
6	33	Turkey	Europe & Central Asia	0.88	17.54	10.53	33.33	34.21	3.51	114	0.14	0.16
7	36	Malaysia	East Asia & Pacific	1.82	8.18	1.82	35.45	50.00	2.73	110	0.33	0.33
8	40	Argentina	Latin America & Caribbean	2.41	7.23	3.61	32.53	49.40	4.62	83	0.18	0.22
9	41	Colombia	Latin America & Caribbean	-	13.70	5.48	19.18	57.53	4.11	73	0.14	0.27
10	42	Thailand	East Asia & Pacific	1.39	4.17	1.39	22.22	69.44	1.39	72	0.10	0.14
Remaining upper-middle income countries (42 in total)				1.71	15.74	4.16	25.86	50.14	2.39	300	0.16	0.28
Lower-middle income countries												
Lower-middle income overall				<u>1.58</u>	<u>11.39</u>	<u>4.03</u>	<u>16.84</u>	<u>64.72</u>	<u>1.43</u>	<u>1,396</u>	<u>0.05</u>	<u>0.19</u>
1	15	India	South Asia	0.30	7.88	3.54	17.03	67.22	4.82	1,016	0.07	0.38
2	43	Ukraine	Europe & Central Asia	1.61	9.68	4.84	40.32	40.32	3.23	62	0.14	0.40
3	44	Egypt, Arab Rep.	Middle East & North Africa	3.39	38.88	1.69	6.78	47.46	1.69	59	0.05	0.16
4	47	Morocco	Middle East & North Africa	-	30.77	2.56	30.77	33.33	2.56	39	0.11	0.32
5	55	Philippines	East Asia & Pacific	13.79	-	-	51.72	31.03	3.45	29	0.03	0.08
6	62	Vietnam	East Asia & Pacific	0.00	19.05	9.52	14.29	57.14	-	21	0.02	0.06
7	63	Sri Lanka	South Asia	5.26	-	5.26	21.05	68.42	-	19	0.09	0.22
8	68	Tunisia	Middle East & North Africa	-	5.88	11.75	17.65	64.71	1.77	17	0.14	0.40
9	70	Ghana	Sub-Saharan Africa	-	-	-	13.33	80.00	6.67	15	0.05	0.21
10	72	Bangladesh	South Asia	-	-	6.87	13.33	80.00	-	15	0.01	0.04
Remaining lower-middle income countries (35 in total)				1.24	11.49	3.81	14.53	67.82	1.11	164	0.04	0.17
Low-income countries												
Low-income overall				-	<u>5.97</u>	-	<u>19.61</u>	<u>75.63</u>	-	<u>53</u>	<u>0.02</u>	<u>0.35</u>
1	85	Sudan	Sub-Saharan Africa	-	27.27	-	-	72.73	-	11	0.02	0.41
2	89	Korea, DPR.	East Asia & Pacific	-	11.11	-	11.11	77.78	-	9	0.03	n/a
3	111	Madagascar	Sub-Saharan Africa	-	25.00	-	50.00	25.00	-	4	0.01	0.31
4	112	Ethiopia	Sub-Saharan Africa	-	-	-	100.00	-	-	4	0.00	0.04
5	113	Gambia, The	Sub-Saharan Africa	-	-	-	-	100.00	-	4	0.18	2.21
6	120	Somalia	Sub-Saharan Africa	-	-	-	-	100.00	-	3	0.02	0.44
7	123	Central African Republic	Sub-Saharan Africa	-	-	-	-	100.00	-	2	0.04	0.86
8	126	Syrian Arab Republic	Middle East & North Africa	-	50.00	-	-	50.00	-	2	0.01	0.18
9	127	Uganda	Sub-Saharan Africa	-	-	-	50.00	50.00	-	2	0.00	0.05
10	128	Eritrea	Sub-Saharan Africa	-	-	-	100.00	-	-	2	0.06	n/a
Remaining low-income countries (19 in total)				-	-	-	16.87	83.33	-	10	0.01	0.16

Notes: Underlined values depict values that largely differ from country-group average.

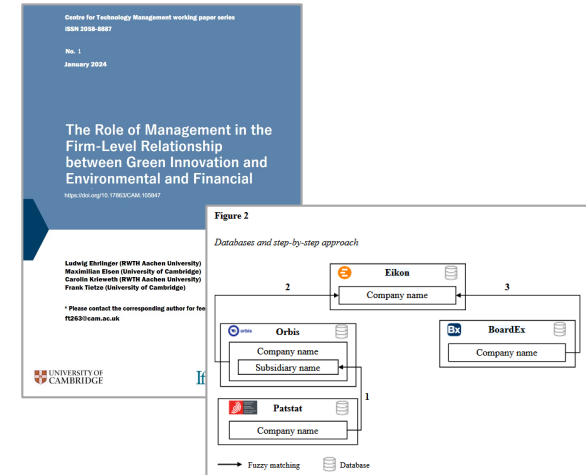
¹ Population calculation is based on high-value adaptation inventions per 100,000 people; population data was retrieved from World Bank for the year 2020.

² GDP Calculation is based on high-value adaptation inventions per 1 billion GDP; GDP data was retrieved from World Bank for the year 2020.

MATCHING PATENT DATA WITH OTHER DATASETS REMAINS CHALLENGING



- Classical approach in econ studies:
NBER dataset that matches US publicly listed firms with patent data from the USPTO between 1980 and 2015 ¹
- Text based matching approaches:
 - Levenshtein ratios of string similarity (Edit Distance) ²
 - Jaro Winkler distance ³
 - Python fuzzy-matching algorithm 'FuzzyWuzzy' using term frequency-inverse document frequency (TF-IDF) ⁴



“This matching process accounts for a significant part of the work for this study since it is both time-consuming and fundamental for having an insightful dataset.” ⁴

¹ Arora, A., Belenzon, S., & Sheer, L. (2021). Matching patents to compustat firms, 1980–2015: Dynamic reassignment, name changes, and ownership structures. *Research Policy*, 50(5), 104217.

² Levenshtein, Vladimir I. (February 1966). "Binary codes capable of correcting deletions, insertions, and reversals". *Soviet Physics Doklady*. 10 (8): 707–710. Example: e.g. Neuhäusler, Peter, Rainer Frietsch, Carolin Mund, and Verena Eckl. (2016) 'Identifying the Technology Profiles of R&D Performing Firms — A Matching of R&D and Patent Data'. *International Journal of Innovation and Technology Management*

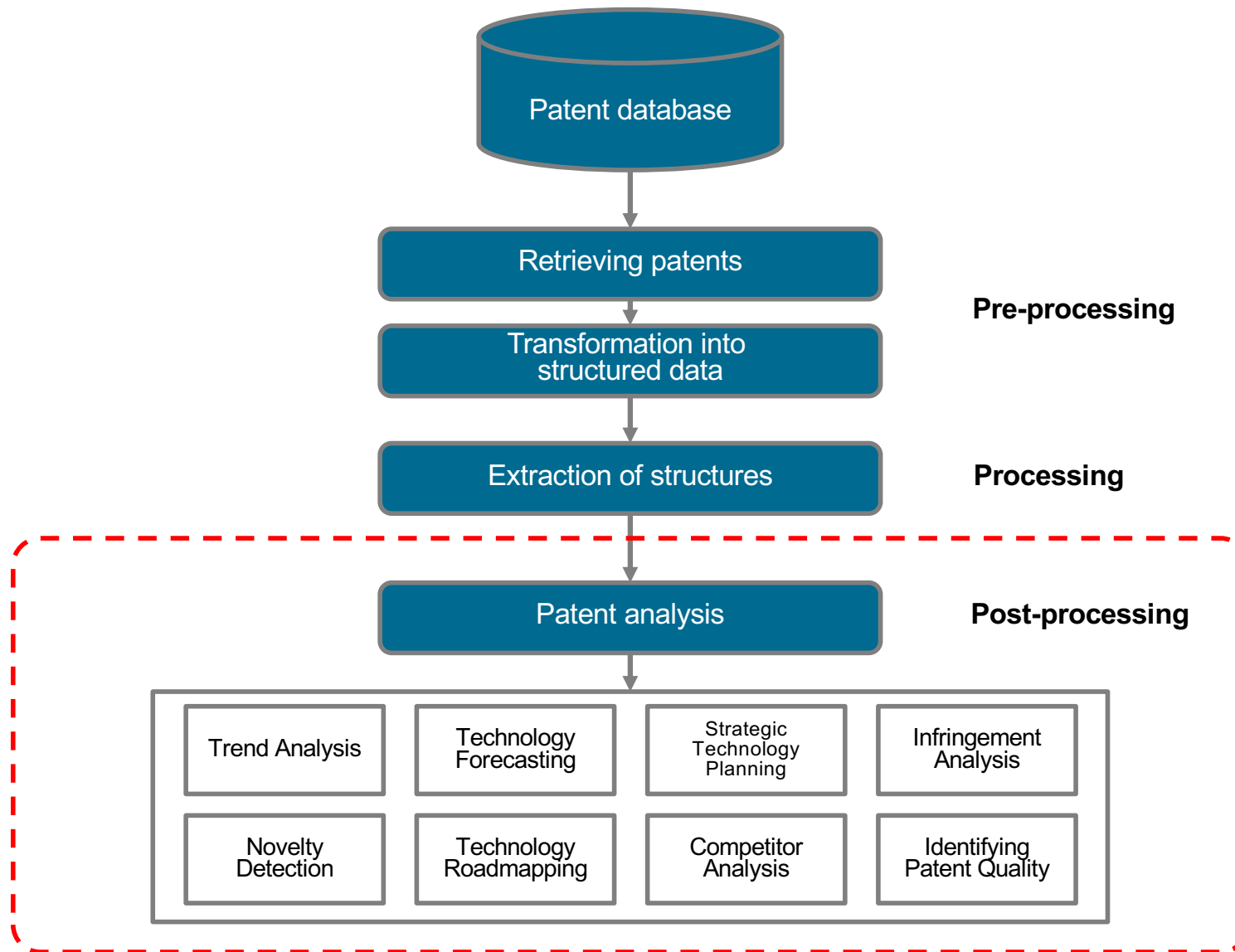
³ Jaro, Matthew A. (1 June 1989). "Advances in Record-Linkage Methodology as Applied to Matching the 1985 Census of Tampa, Florida". *Journal of the American Statistical Association*. pp. 414–420.

⁴ Ehrlinger, Ludwig, Maximilian Elsen, Carolin Krieweth, and Frank Tietze. (2024) 'The Role of Management in the Firm-Level Relationship between Green Innovation and Environmental and Financial Performance'. <https://www.repository.cam.ac.uk/handle/1810/364028>.

See further: 'String Similarity Metrics – Edit Distance | Baeldung on Computer Science', 14 November 2020. <https://www.baeldung.com/cs/string-similarity-edit-distance/>



THE PATENT ANALYSIS WORK FLOW





THE PATENT INDICATOR “JUNGLE”: NUMEROUS (TOO MANY?) PATENT INDICATORS HAVE BEEN DEVELOPED

> 1,000 citations (Google Scholar)

Patent indicator	Definition	Meaning
Patent activity (PA_{iF})	Patent applications (PA) of firm i in technological field (TF) F	Extent of R&D expenditures of firm i in TF F (interest of firm i in TF F)
Technology share (based on patent applications)	PA_{iF}/PA of all competitors in TF F	Competitive technological position of firm i in TF F (quantitative)
R&D emphasis	PA_{iF} /Number of firm's (i) total patent applications	Importance of technological field F for firm i (R&D emphasis)
Co-operation intensity	Number of joint patent applications with partners in TF F/PA_{iF}	Access of firm i to external knowledge (and identification of partners)
Share of granted patents (Q_i)	Granted patents of firm i in TF F/PA_{iF}	Technological quality of firm i 's patent applications
Technological scope (Q_2)	Diversity and number of IPC classes in firm i 's patent applications (PA_{iF})	Technological quality of firm i 's patent applications
International scope (Q_3)	Size of patent family and share of triad (US, JP and EPO) patents of PA_{iF}	Economic quality of firm i 's patent applications
Citation frequency (Q_4)	Average citation frequency of PA_{iF}	Economic quality of firm i 's patent applications
Average patent quality (PQ_{iF})	Sum of all indicators of patent quality (Q_1-Q_4)	Average total quality of all patent applications of firm i in TF F
Patent strength (PS_{iF})	Product of average patent quality (PQ_{iF}) and patent activity (PA_{iF})	Technological strength of firm i in TF F
Technology share (based on patent strength)	PS_{iF}/PS of all competitors in TF F	Competitive technological position of firm i in TF F (qualitative)
Relative technology share	$PS_{iF}/\text{Max. patent strength of a firm in TF } F$	Distance of firm i to the technological leader in TF F

Ernst, H. (2003). Patent information for strategic technology management. World Patent Information, 25(3), 233-242.

See also:

Lanjouw, J. O. & Schankerman, M. Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators. The Economic Journal 114, 441–465.

OECD Science, Technology and Industry Working Papers. Measuring Patent Quality: Indicators of Technological and Economic Value, 2013.

Table 1

Citation-based measures - patent level.

Measure	Description	Source	Used by	Equation
Citation Index (CI)	CI is the count of the citations received by a company's patents from subsequent patents. It is used to evaluate the technological impact of patents. The number of forward citations mirrors the technological importance of the patent for the development of subsequent technologies, and also reflects the economic value of inventions [54,54]. Patents with high values of CI are often important inventions or fundamental to future inventions, making these inventions valuable [54] due to the cumulative nature of the process. FCF is defined as the number of forward citations received by a patent per year. This is an indication of the impact of a company's patents. FCF can only be compared within a technological area for a particular year, since the number of citations changes per year.	[10,11,15,23,25,34,54,56,57]	[24,37,41,43,50,53,58]	$CI = \sum_{t=1}^n \sum_{i=1}^n \text{Citations citing patent}_i$ where t is the year of patent publication and n is the publication year of subsequent patents citing the original patent.
Forward Citation Frequency (FCF)	The generality of a target patent indicates the diversity of citing patents, i.e. the patents that cite the target patent. The index is defined between zero and one, and the measure is high if subsequent patents belonging to a wide range of fields cite a patent. If most citations are concentrated in a few fields the generality index is low [63]. Note the numerical range of the index with 0-0.44 (low), 0.45-0.65 (mid) and 0.66-1 (high).	[24,33,34,39,59,60]	[27,54,61]	$FCF = \frac{CI}{\text{Patent age}}$
Generality	The influence is defined as the number of forward citations a patent received from subsequent patents in the first years since its publication. The higher influence index suggests that a patent has influenced and impacted the technology scope of subsequent patents on publication.	[25,63]	[13,24,30,39,64]	$\text{Generality} = 1 - \sum_{i=1}^n \left(\frac{\text{Number of Citing Patents}}{\text{Number of Citing Patents}} \right)^2$ where SI is the set of classes of citing patents
Influence		[6]	[6,63]	$\text{Influence}_i = 1 + \left(\sum_{t=1}^n \text{CI}_i \right)$ $\text{CI}_i(t)$ is the number of forward citations the i patent received from following patents up to five years after the publication date, where t is the year.

Table 2

Citation-based measures - portfolio level.

Measure	Description	Source	Used by	Equation
Current Impact Index (CII)	CII is the number of times a company's previous 5 years of patents are cited in the current year, relative to all patents in the U.S. patent system. It can measure the influence of a company in the last 5 years, and indicates patent portfolio quality. CII is a synchronous indicator, which looks back from the current year to the previous five years. In general, patents with higher CII values often represent stronger technological ability. It has the advantage to foresee the development of a technology.	[66]	[41,51,52,54,67]	$CII = \frac{\sum_{i=1}^n \text{Citations}_i}{\sum_{i=1}^n \text{Patents}_i}$ CII represents the number of times a patent of company i has been cited in a certain year, from previous 5 years. K_i is the number of patents, company i produced the past 5 years.
Herfindal-Hirschman Index of Patents	HHI describes the concentration of patents across patent classes, and is used to measure the concentration level of a firm's technological capability. HHI is described as a patent quality indicator and is used to explore the relationship between patent quality and market value. According to [68], a HHI index constructed from a small number of counts will generally be biased downwards. To avoid the bias, we also use the HHI index adjusted for bias measure [68].	[24,25,47,68]	[24,34,39,47,60]	For a set of N patents falling into n classes, with N_i patents in each class ($N_i \geq 0, i = 1 \dots n$), the HHI is calculated $HHI \text{ of patents} = \sum_{i=1}^n \left(\frac{N_i}{N} \right)^2, 0 \leq HHI \text{ of patents} \leq 1$ Adjusted bias measure $\eta = \frac{N \cdot HHI - 1}{N - 1}$
Hindrance Index (HI)	The measure of hindrance, of the assignee of the target patent is defined as the ratio of "X" and "Y" citations received from competitors to all "X" and "Y" citations received. If the firm received "X" or "Y" references from future patents they can be seen to hinder or block subsequent patents.	[13]	[13]	$HI = \frac{\sum_{i=1}^n \text{X and Y forward citations from competitors}}{\sum_{i=1}^n \text{X and Y forward citations received}}$ Forward citations are split into "citations by competitors (the no. of patents with a different assignee than the focus patents)" and "total citations by all other non-competitive companies (the no. of patents developed conjointly between competitors)".
Relative Patent Position (RPP)	RPP of a company in its most important technological field means the patent counts owned by the company in its technological field where it has more patents than in others divided by the patent counts of the leader in the technological field. RPP is used to measure the degree of leading of the company in the technological field, the higher the value, the more leading position the assignee of that patent has in the field.	[69]	[26,52,70]	$RPP = \frac{\sum_{i=1}^n \text{Patents owned by the company}}{\sum_{i=1}^n \text{Patents owned by the leader}}$ The leader in a technological field is defined as the company with the largest amount of patents granted in the field.
Technology Strength (TS)	TS indicator is used to measure the scale of a company in a specific technological field, indicating patent portfolio strength.	[54,71]	[51,54,55,67]	$TS = \text{Number of Patents} \times CII$ It makes the assumption that the number of patents of a company in a specific field is of similar quality and impact as the company's recent patents.

Table 1 in Aristodemou, L. and F. Tietze (2018). "Citations as a measure of technological impact: A review of forward citation-based measures." World Patent Information 53: 39-44.

REMAINING AMBIGUITY IN PATENT FAMILY DEFINITIONS

Simple family: n=3

Simple Family

A simple patent family is a group of patent documents that stem from the same initial document, called the priority document. For example, an applicant might file a patent application in one country, then file other applications in other countries. The simple family covers one single invention.¹

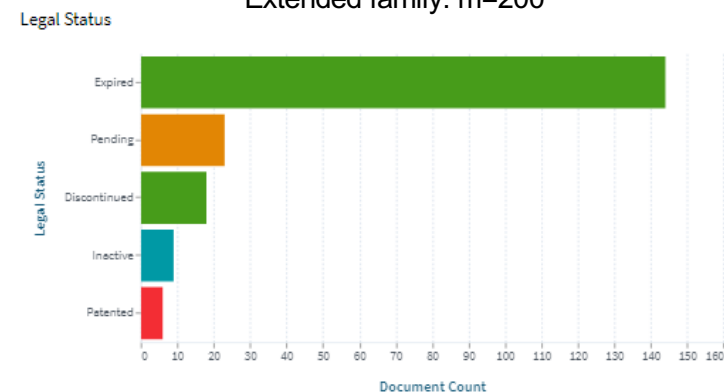
The screenshot shows a search results page for patent families. It lists three entries, each representing a different patent family for the same invention. Each entry includes the patent number, status, and a brief description of the invention.

Patent Number	Status	Description
GR 2021181 T3	Granted Patent	Family: 3s / 2020s: Family Jurisdictions: GR, US, ES Legal Status: Expired
ES 2160901 T3	Granted Patent	Family: 3s / 2020s: Family Jurisdictions: GR, US, ES Legal Status: Expired
US 596332 A	Granted Patent	Family: 3s / 2020s: Family Jurisdictions: GR, US, ES Legal Status: Expired

Extended Family

An extended patent family is a collection of patent documents covering a technology – more than one single invention. The technical content covered by the applications is similar, but not necessarily the same. Members of an extended patent family will have at least one priority in common with at least one other member - either directly or indirectly.²

Extended family: m=200



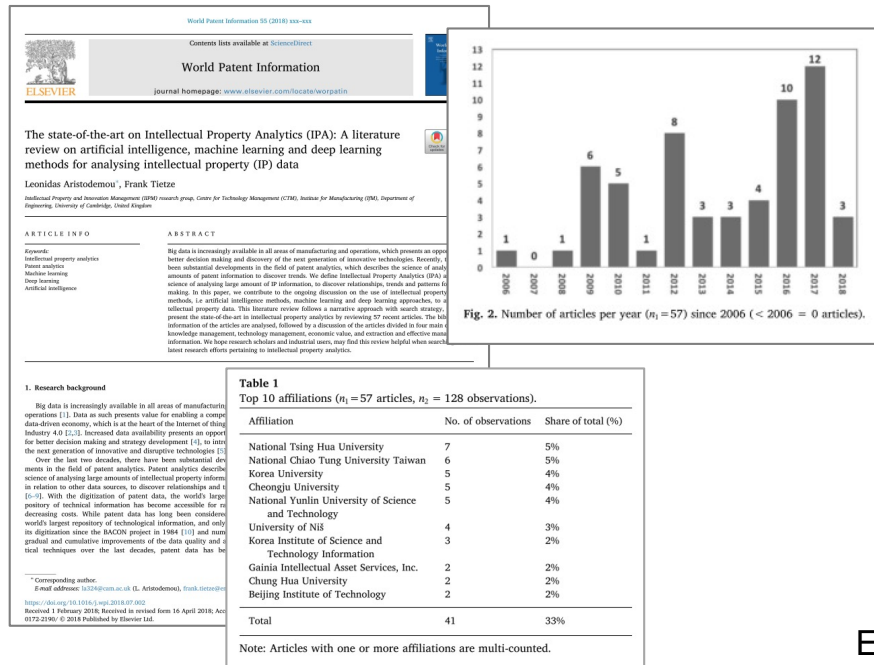
¹ See also: <https://www.epo.org/en/searching-for-patents/helpful-resources/first-time-here/patent-families/docdb>

² See also: <https://www.epo.org/en/searching-for-patents/helpful-resources/first-time-here/patent-families/inpadoc>

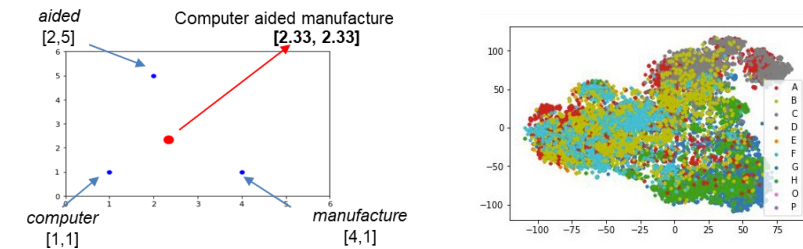
NEED FOR EVEN MORE GUIDANCE AND STANDARDIZATION



INCREASINGLY STUDIES USE AI ALSO FOR UNSTRUCTURED TEXT DATA



Text data preparation remains a challenge



First studies experiment with LLMs to analyse patent text ¹

Example from current own work: Text-based classifications of Patent Litigation cases using OpenAi GPT-4o model

Approach	Method	Authors	Approach	Method	Authors
Artificial Neural Networks (ANN)	Back Propagation learning (BP)	[30,31,36,39-51]	Regression	Linear	[13,30,37,54]
	Evolutionary sigmoidal unit	[52]		Logistic	[62,68,69]
	Evolutionary product unit	[53,54]		Conditional random fields (CRF)	[29,34,58,70]
	Extreme learning machine (ELM)	[43,47,55]		Latent Dirichlet Allocation (LDA)	[56,71]
Clustering	Growing cell structure, paired with Girvan-Newman clustering algorithm	[56]	Support Vector Networks (SVN)	Naive Bayes	[62,65]
	Restricted Boltzmann machines	[57]		Hidden Markov Model (HMM)	[72]
	K-means (and derivatives)	[33,35,52,58,59]		Support Vector Clustering (SVC)	[34,38,45,60,73-76]
	Self organising maps (SOM)	[36,39,40,60]		Semantic Support Vector Machine (SSVM)	[70]
Deep Learning (DL)	Deep Belief Networks (DBN)	[57]	Text mining approaches	Dictionary-based approach	[34,58]
	Reinforcement Learning (RL)	[61]		Natural Language Processing (NLP)	[34,68]
	Boosting	[29]		Rule-based approach	[34,62]
	Random Forest	[63]		Semantic based ontology	[49,70,71]
Decision tree	Stacking	[63]			
	Classification and Regression Tree (CART)	[64,65]			
	C4.5	[62]			
	Linear Discriminant Analysis (LDA)	[50,66]			
Dimensionality Reduction	Multi-dimensional scaling (MDS)	[67]			
	Principal Component Analysis (PCA)	[31,33,54]			
	Quadratic Discriminant Analysis (QDA)	[50]			
	Singular Value Decomposition (SVD)	[33]			



THE PATENT ANALYSIS WORKFLOW IS CHANGING

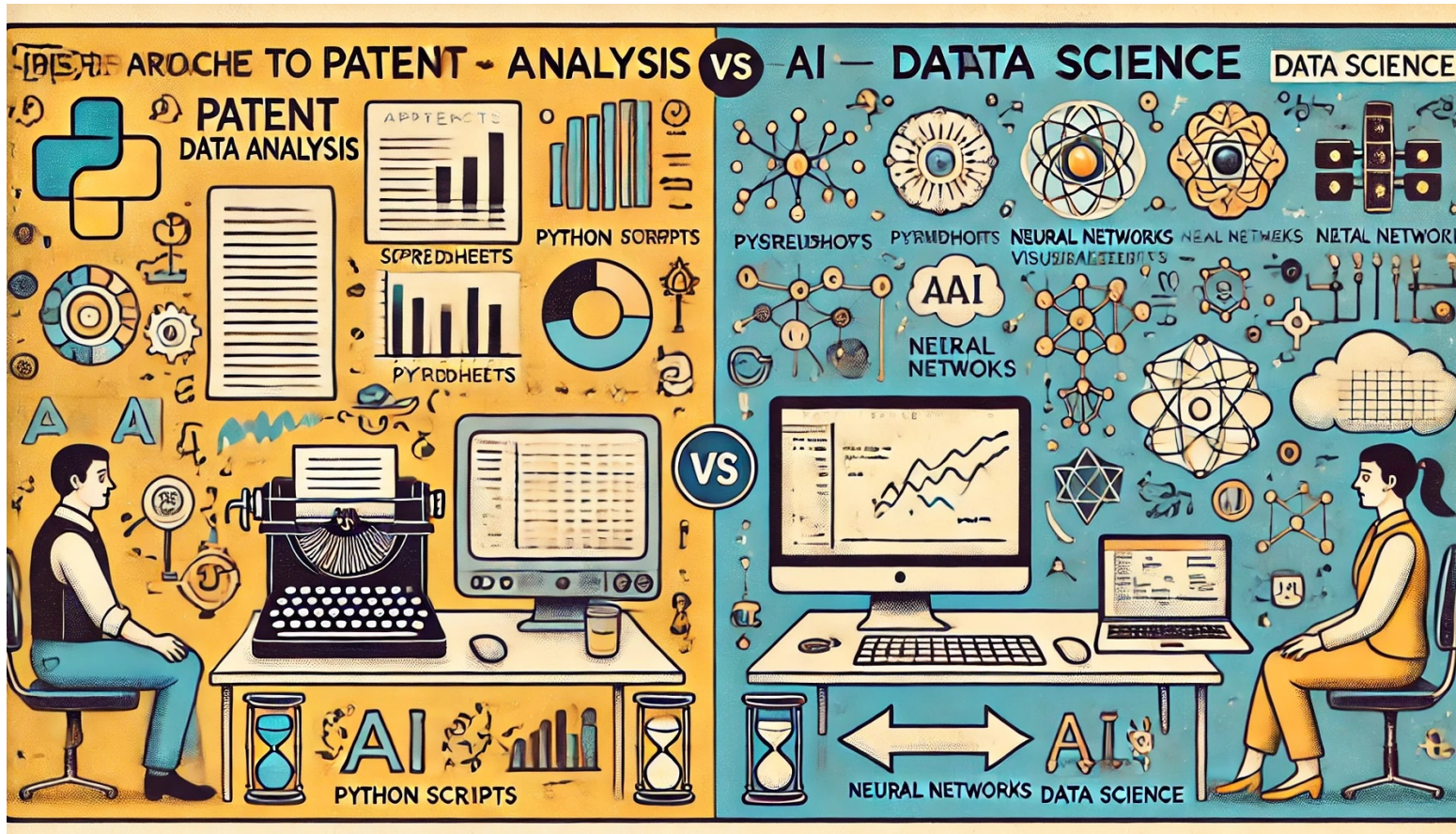


Image generated with <https://chatgpt.com>

Deploying AI for patent analysis requires new data science skillsets and tools

SOME ADDITIONAL CONSIDERATIONS: WHAT ELSE FOR COP?

- Raising the acceptance of patent studies by showcasing the impact of patent studies, e.g. collecting impact case studies
- Empowering policy makers to better understand patent data, e.g. training
- To what extent to engage with related communities, e.g. academics to address technical challenges, such as indicators/metrics, dataset matching, automated text analysis



Image generated with <https://chatgpt.com>

CONCLUSIONS: WHAT COULD THE COP DO?

- Defining its remit and positioning of the COP, i.e. what scope of studies to consider getting involved
- Work together to increase quality of “raw” data, collaborate with academic researchers
- Collaborate on establishing best practices, e.g. for analysis, indicator usage, dataset matching, but also visualisation and dataset sharing
- Help CoP members to develop data science/AI skill sets for analysing data, e.g. develop training material
- Educate policy makers about the use of patent data, e.g. demonstrate impact by showcasing examples
- Work to further reduce access barriers to patent data users, such as in LMIC
- Consider how CoP and patent analytics can best support addressing urgent global challenges, such as climate change, SDGs (e.g. improved Y02 classification)



Image generated with <https://chatgpt.com>





THANK YOU

Prof. Dr. Frank Tietze

Innovation and IP Management (IIPM) Laboratory

www.iipm.eng.cam.ac.uk

frank.tietze@eng.cam.ac.uk

Department of Engineering

Institute for Manufacturing (IfM)

Centre for Technology Management

17 Charles Babbage Road

Cambridge, CB3 0FS

United Kingdom

