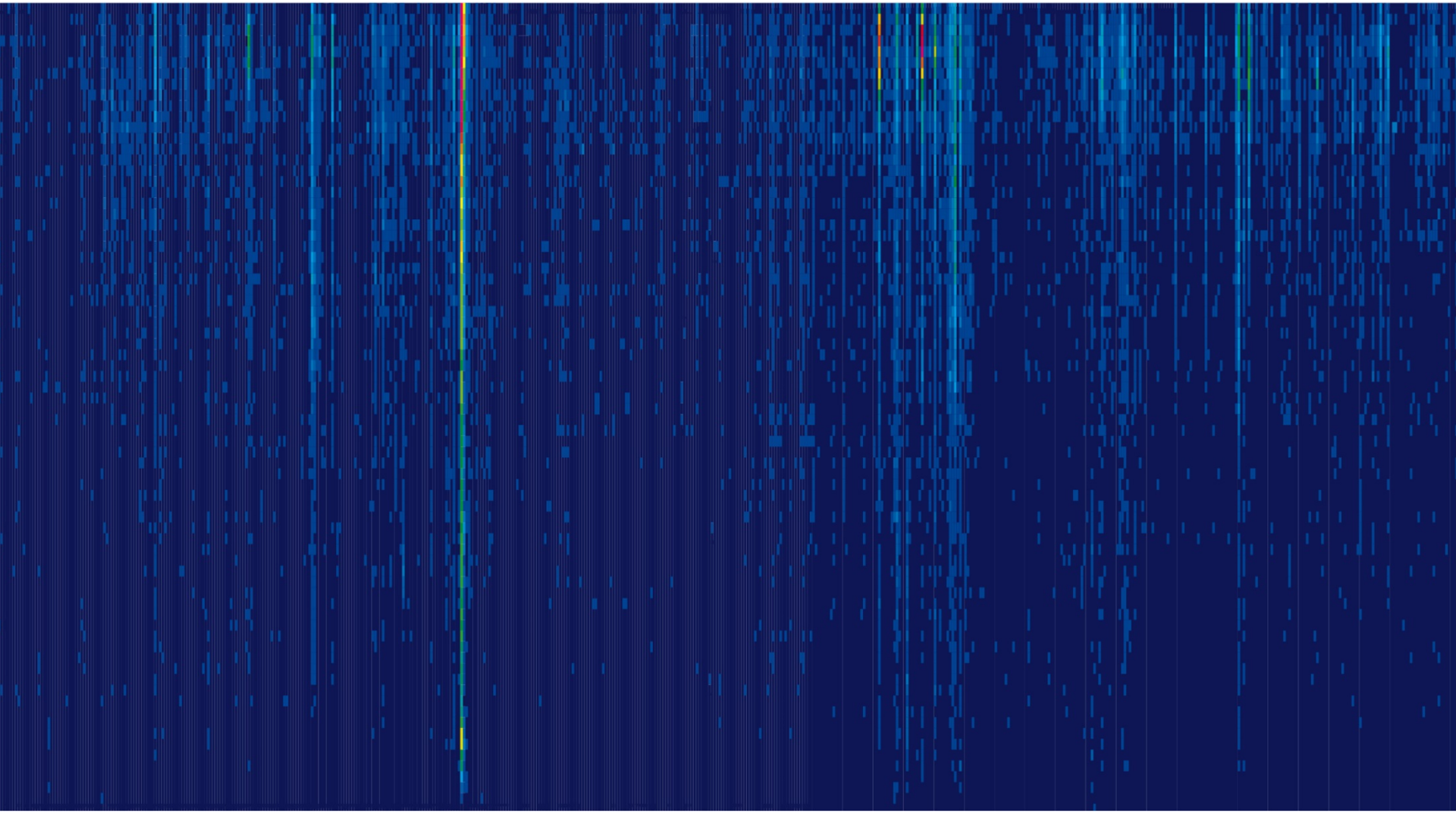
Europäisches Patentamt
European Patent Office
Office européen des brevets

Know what you're analysing

Using patent statistics in a business context



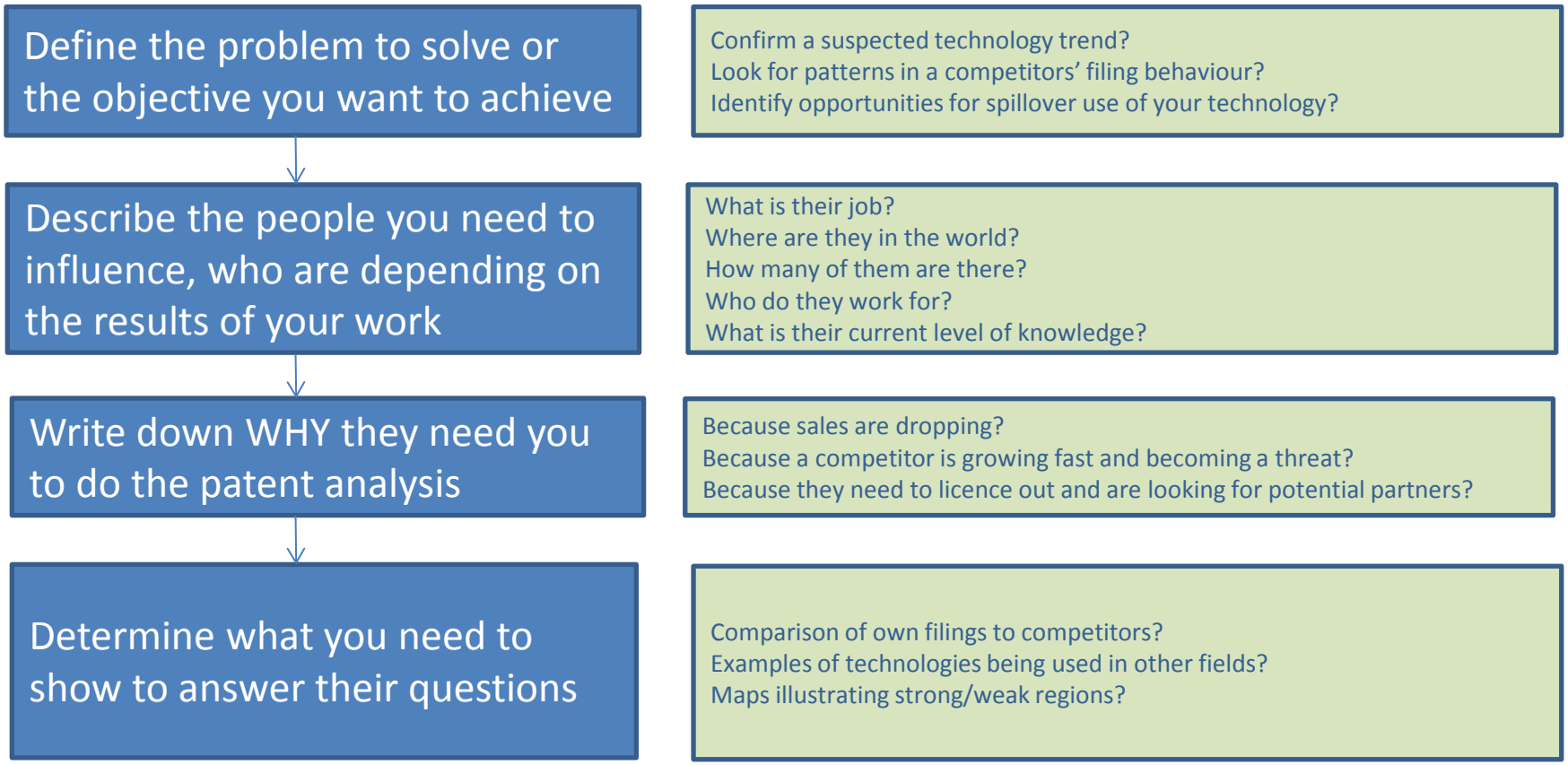


It's not about finding a needle in a haystack

It's about helping someone to take decisions

- **understand the objective**
- **understand the data**
- **then do your analysis and present your results**

Patent analysis



If you don't understand the objective

And you don't know WHY you're doing the analysis

... you will never be able to decide which data to use

Traditional patent statistics and data/text mining: important data issues

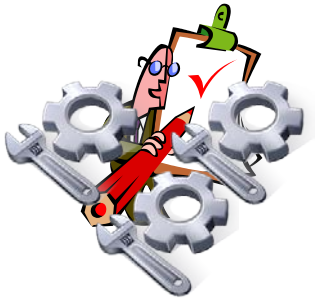
- Data coverage (e.g, with respect to time, patent authorities)
- Completeness
- Harmonisation of data
- Accessibility of data

over 90 offices

Bibliographic data
Citations
Full text
Images
Legal Status

raw data

process data

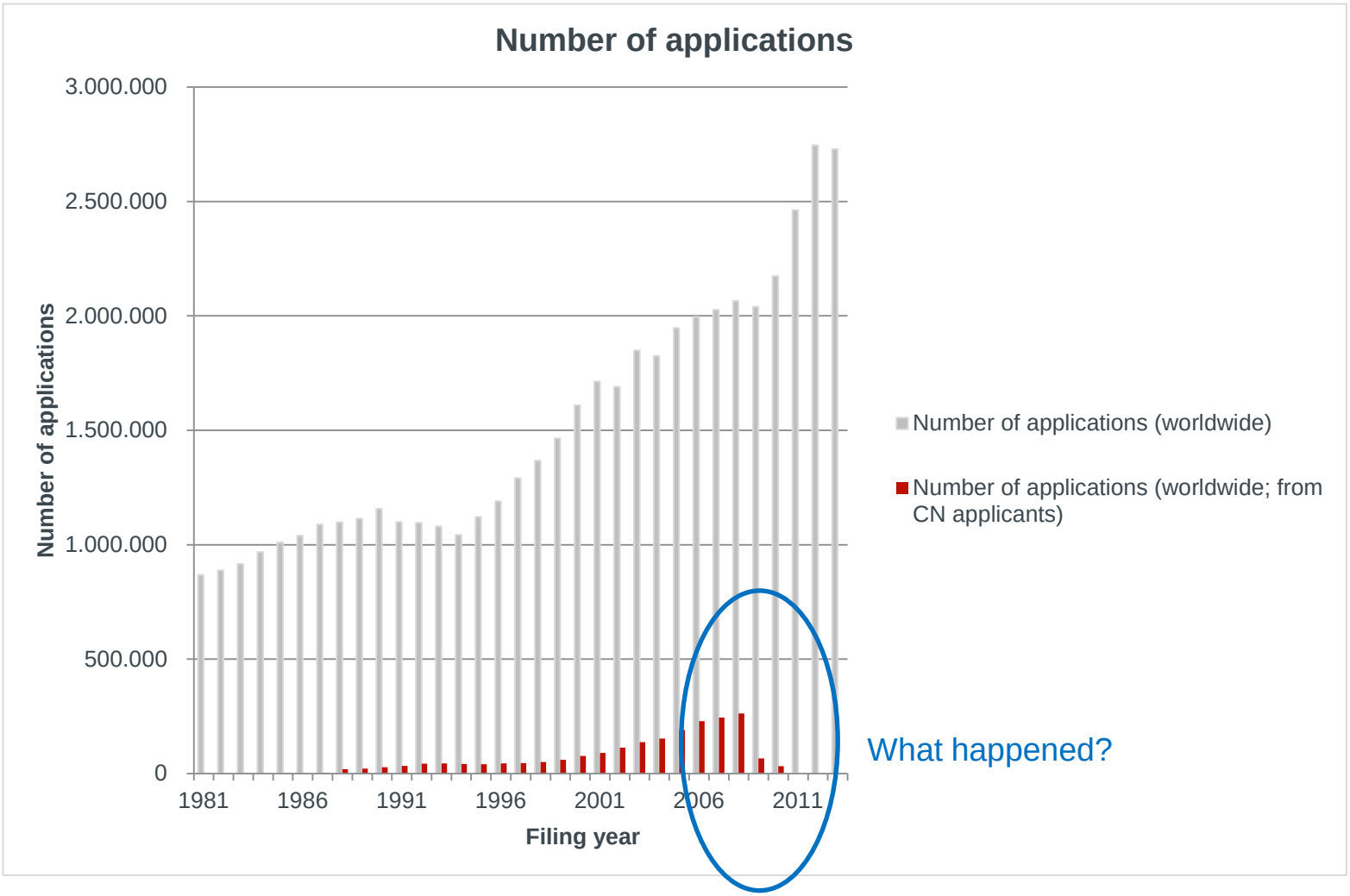


valid data

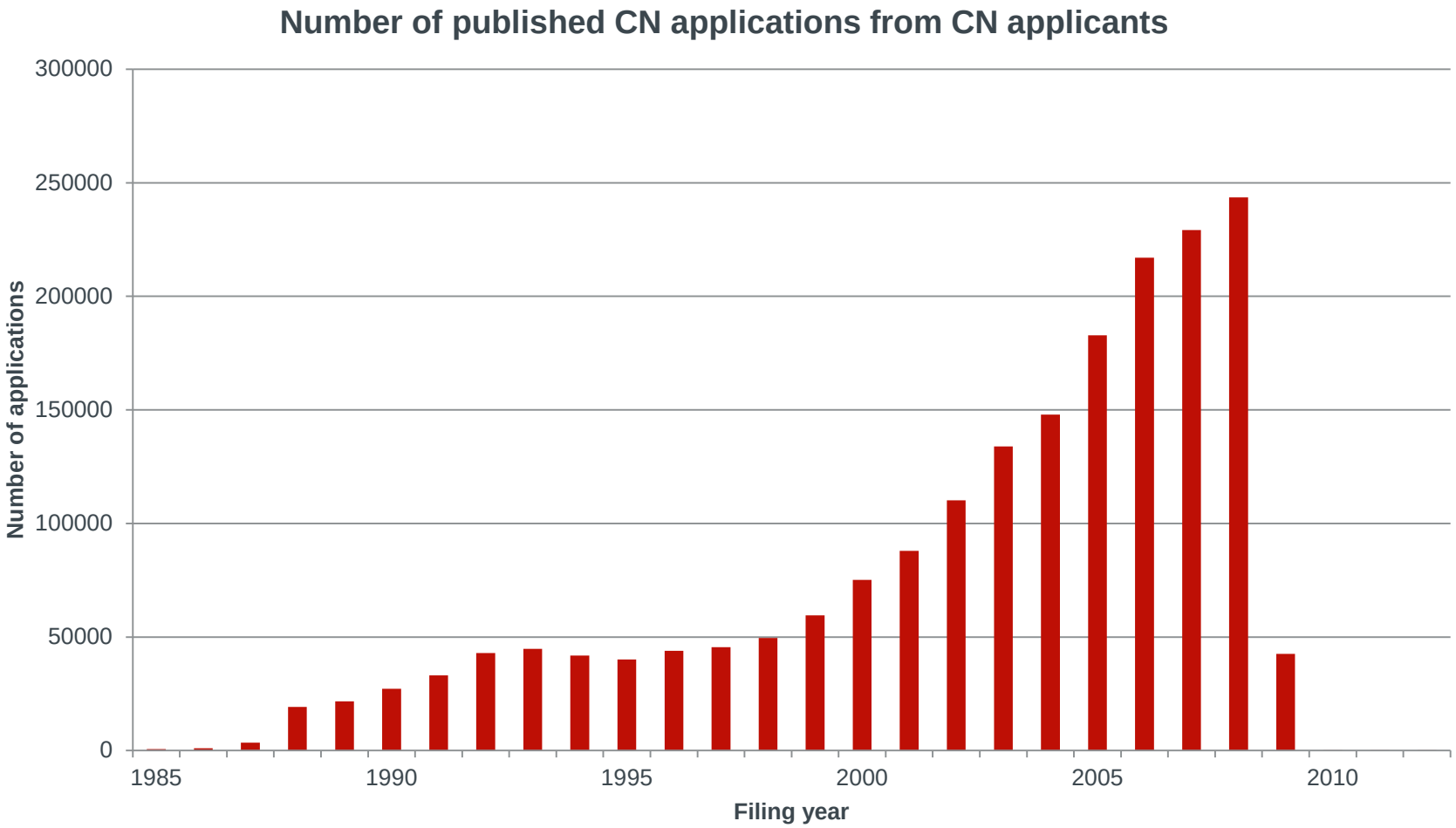
EPO Master Databases

services & products

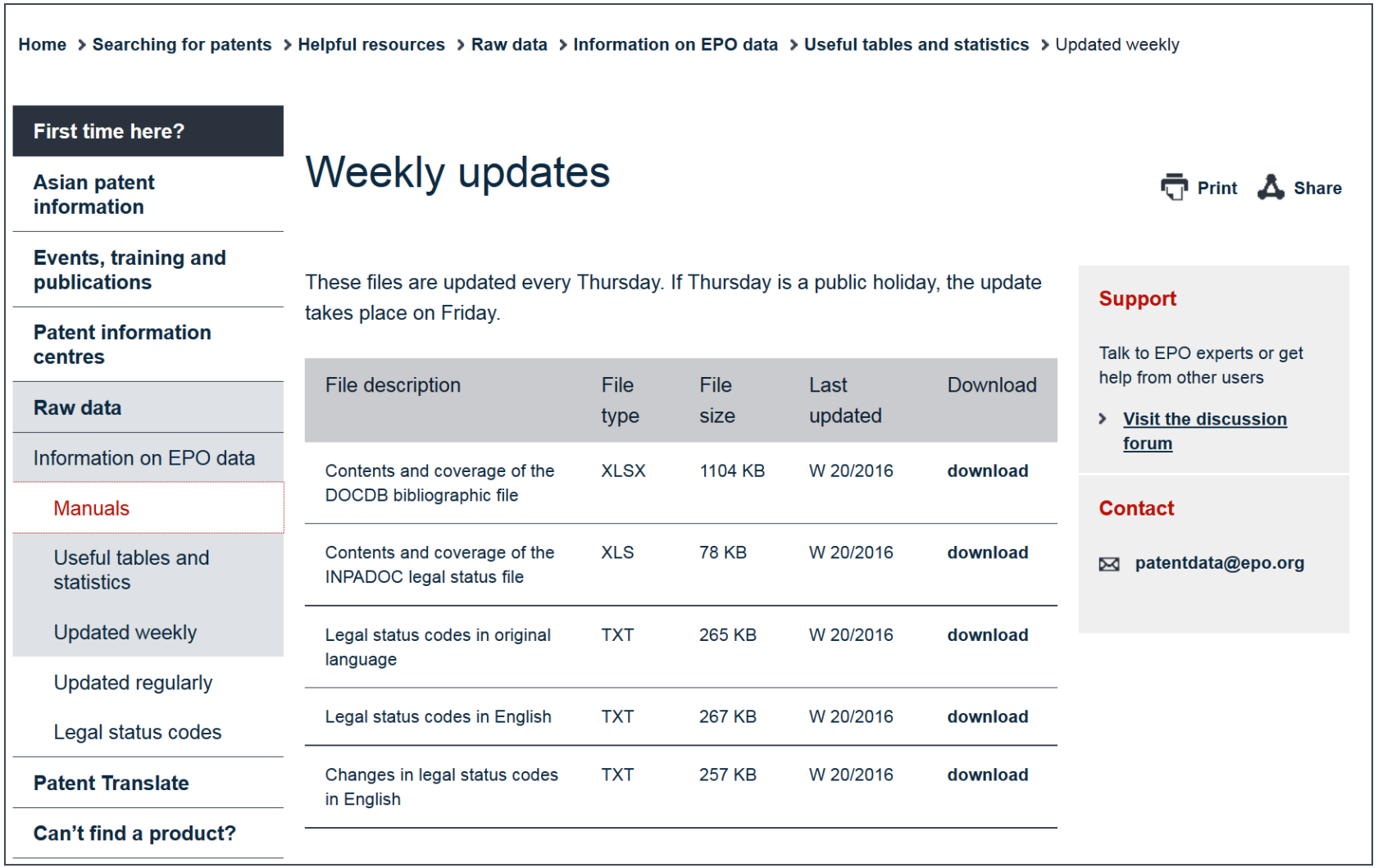
Sometimes the appearances can be deceiving ...



Sometimes the appearances can be deceiving ...



Coverage information on the EPO website



Home › Searching for patents › Helpful resources › Raw data › Information on EPO data › Useful tables and statistics › Updated weekly

First time here?
Asian patent information
Events, training and publications
Patent information centres
Raw data
Information on EPO data
Manuals
Useful tables and statistics
Updated weekly
Updated regularly
Legal status codes
Patent Translate
Can't find a product?

Weekly updates

Print Share

These files are updated every Thursday. If Thursday is a public holiday, the update takes place on Friday.

File description	File type	File size	Last updated	Download
Contents and coverage of the DOCDB bibliographic file	XLSX	1104 KB	W 20/2016	download
Contents and coverage of the INPADOC legal status file	XLS	78 KB	W 20/2016	download
Legal status codes in original language	TXT	265 KB	W 20/2016	download
Legal status codes in English	TXT	267 KB	W 20/2016	download
Changes in legal status codes in English	TXT	257 KB	W 20/2016	download

Support

Talk to EPO experts or get help from other users

› Visit the discussion forum

Contact

patentdata@epo.org

www.epo.org/searching-for-patents/helpful-resources/raw-data/data/tables/weekly.html

Patent statistics: name harmonisation: EPO standardised name table

- Control for variations of a given applicant name
- Variants are mapped to a standard name:

Applicants Variation name
AKTIENGESELLSCHAFT VOLKSWAGEN
FUORUKUSUAAGENUERUKU AG
FUORUKUSUWAAGEN AG
V O L K S W A G E N AKTIENGESE
V W AG
VOLKSWAGEN
VOLKSWAGEN A G
VOLKSWAGEN AG
VOLKSWAGEN AG VW
VOLKSWAGEN AKTIENGESELLSCHAFT
VW
VW AG
VW WOLFSBURG
WOLFSBURG VW
BRASI S A VOLKSWAGEN DO
BRASIL S A VOLKSWAGEN DO
BRASIL SA VOLKSWAGEN

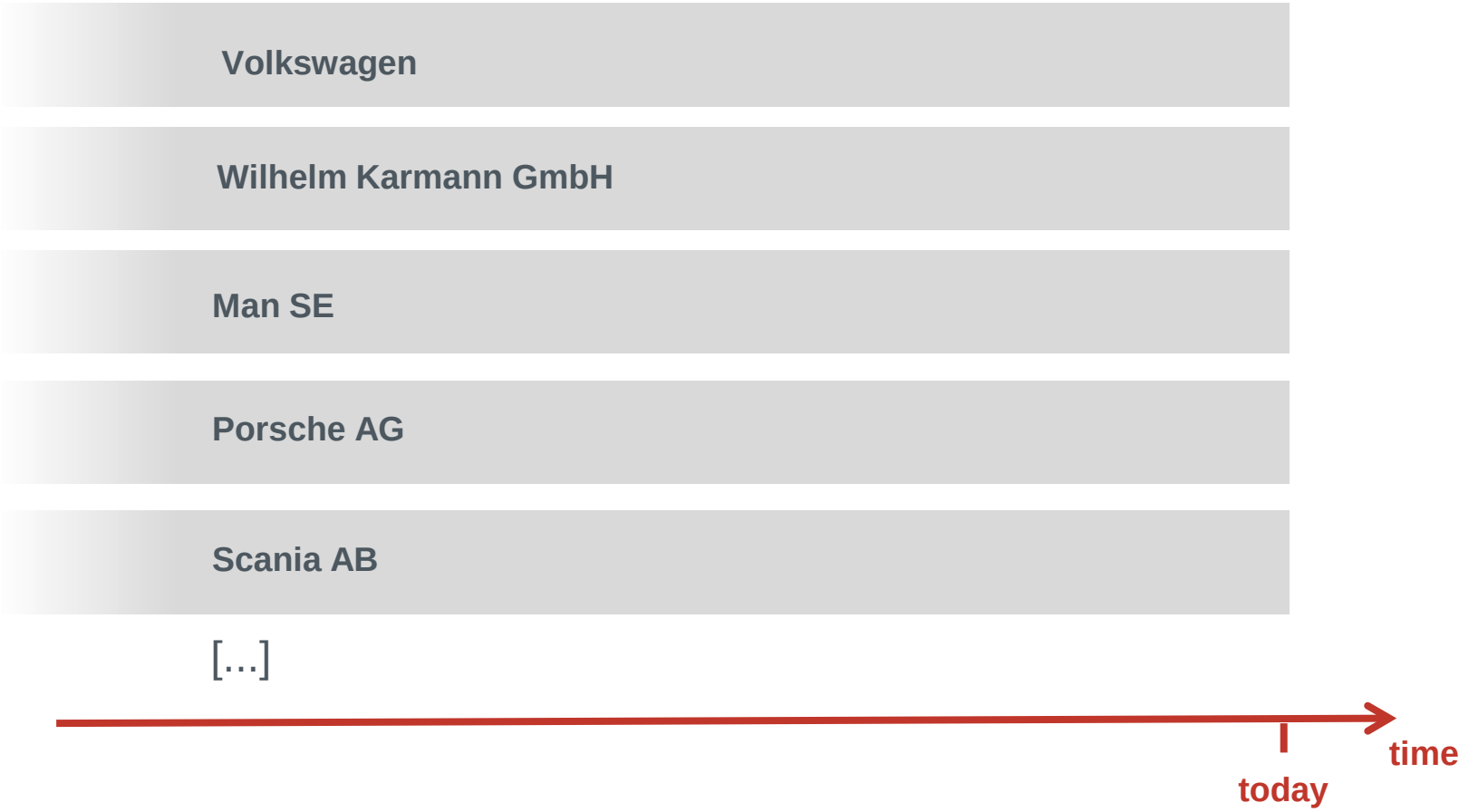
Name harmonisation: corporate structure

- e.g., corporate tree of Volkswagen in Thomson Innovation:

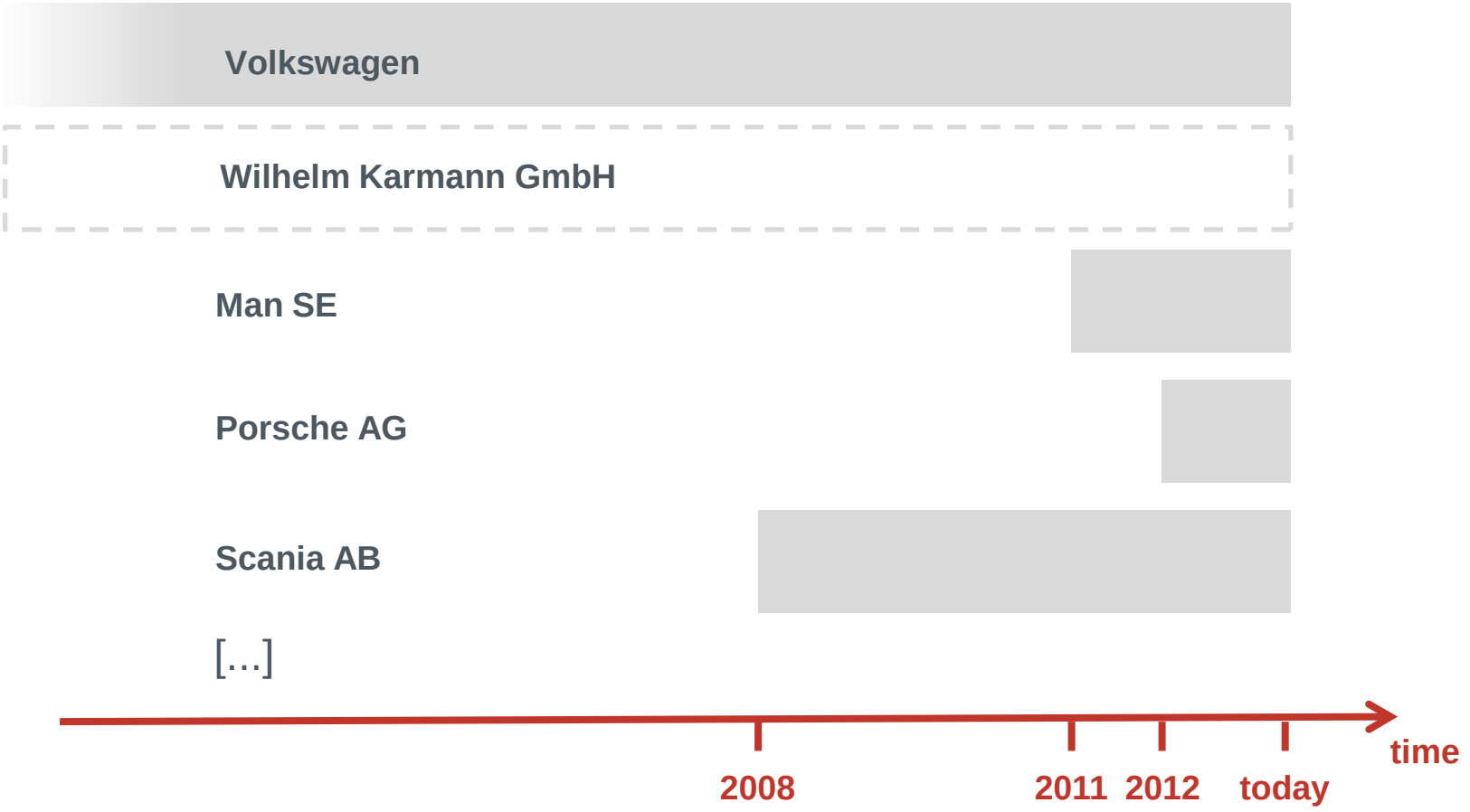
Volkswagen	VOLKSWAGE AG
	VOLKSWAGEN DO BRASIL S/A
	VOLKSWAGEN AG
	VOLKSWAGEN AKTIENEGESELLSCHAFT
	AUDI AG
	SEAT SA
	SKODA AUTO
	[...]
Wilhelm Karmann GmbH	KARMANN WILHELM GMBH
	WILHELM KARMANN GMBH
	WILHEM KARMANN GMBH
	WILHELM KARMANN GMBH I I
	WILHEIM KARMANN GMBH
	WILHELM KARMANN GMBH I L
Man SE	MAN FINANCIAL INC
	MAN NUTZFAHZEUGE OSTERREICH AG
	MAN DIESEL & TURBO SE
	MAN TURBOMASCHINEN AG
	MAN NUTZFAHRZEUGE AKTIENGESELLSHAFT
	MAN GUTEHOFFNUNGSHUTTE AG
	MAN GUTEHOFFNUNGSHUTE AKTIENGESELLSCHAFT
	[...]
Porsche AG	PORSCHE LIZENS-UND HANDELSGESELLSCHAFT MBH & CO KG
	DR ING H C F PORSCHE AG
	F PORSCHE AKTIENGESELLSCHAFT
	DR ING H C F PORSHE AG
	[...]
Scania AB	SCANIA CV AKTIEBOLAC
	SCANIA CV AB (PUB)
	AKTIEBOLAGET SCANIA-VABIS SODERTALJE
	DYNAMATE AB
	[...]

Source: Thomson Innovation, 1790 Analytics

Name harmonisation: Dynamic corporate structure

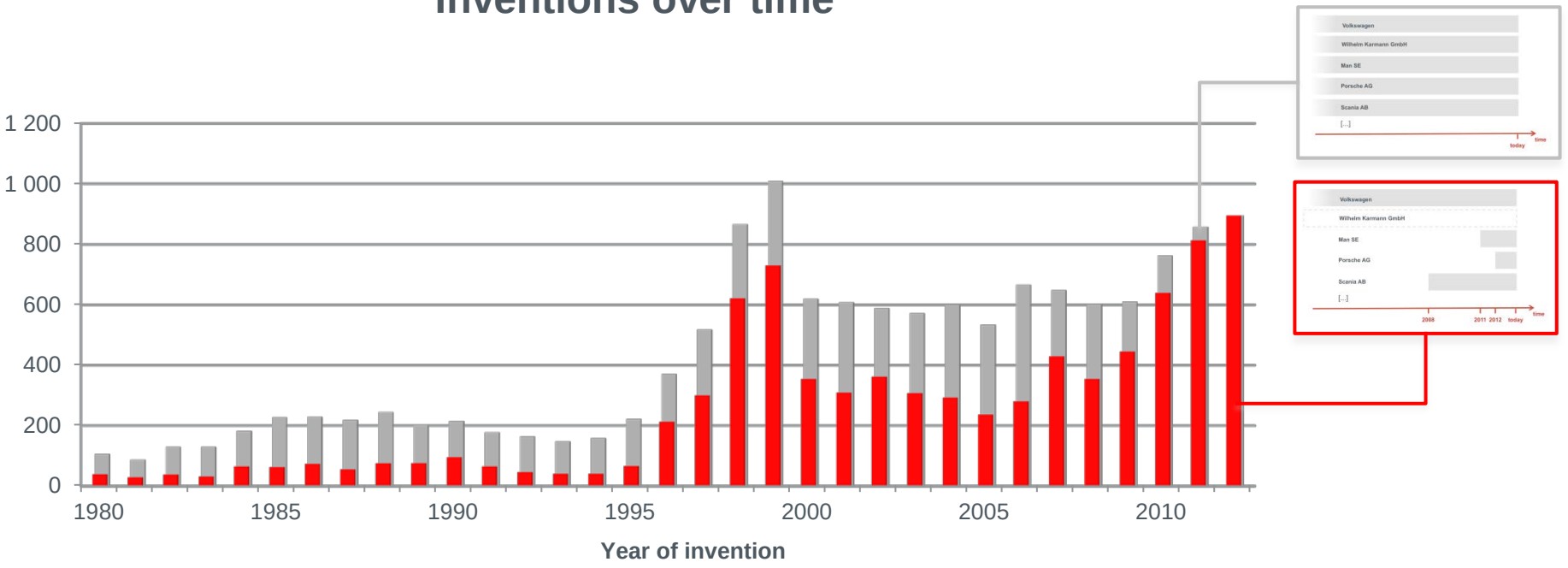


Name harmonisation: Dynamic corporate structure



Name harmonisation: Dynamic corporate structure

Volkswagen: Static vs. Dynamic corporate structure: Inventions over time



Before and after

Increase in average number of name variants (person_IDs) per name

	All applicants			Only subset of top applicants (treated also with *manual layer*)		
	Before harm	**After harm**	**% increase**	**Before harm**	**After harm**	**% increase**
Whole PATSTAT	**1.31**	**1.65**	**26%**	**2.66**	**120.03**	**4421%**
EPO patents	1.37	1.50	9%	4.97	18.14	265%
USPTO patents	1.28	1.48	16%	2.72	39.52	1353%

Source: Julie Callaert, KU Leuven, presentation at EPO Patent Information Conference 2015, Copenhagen

Before and after

Shifts in applicant rankings

Rank	Original name	Patent count
1	MATSUSHITA ELECTRIC IND CO LTD	354036
2	HITACHI LTD	328830
3	SAMSUNG ELECTRONICS CO., LTD.	316871
4	TOSHIBA CORP	285937
5	CANON INC	252268
6	MITSUBISHI ELECTRIC CORP	238153
7	NEC CORP	221929
8	INTERNATIONAL BUSINESS MACHINES CORPORATION	200123
9	FUJITSU LTD	193309
10	ROBERT BOSCH GMBH	173897
11	SONY CORP	170364
12	LG ELECTRONICS INC.	169280
13	GENERAL ELECTRIC COMPANY	166962
14	SIEMENS AKTIENGESELLSCHAFT	151736
15	RICOH CO LTD	148337

Rank	Harmonised name	Patent count
1	PANASONIC CORPORATION	663727
2	TOSHIBA CORPORATION	550834
3	HITACHI	508279
4	CANON	443213
5	SAMSUNG ELECTRONICS COMPANY	413504
6	NEC CORPORATION	409738
7	MITSUBISHI ELECTRIC CORPORATION	389610
8	SONY CORPORATION	344107
9	SIEMENS	341427
10	FUJITSU	335823
11	IBM (INTERNATIONAL BUSINESS MACHINES CORPORATION)	292069
12	PHILIPS ELECTRONICS	290544
13	GE (GENERAL ELECTRIC COMPANY)	255458
14	ROBERT BOSCH	237968
15	TOYOTA MOTOR CORPORATION	235794

Source: Julie Callaert, KU Leuven, presentation at EPO Patent Information Conference 2015, Copenhagen

Patent statistics:
Development over time: which date?

1. Earliest filing date
2. Application date
3. Publication of application
4. Latest publication
5. Date of grant

Patent statistics: which date?

1. Priority date

☺ closest to invention

☹ public is not aware of invention

☹ not available for recent years → distortion

Recommended use:

- proxy for innovation activity
  e.g. evaluation of impact of new policies
- preferred date for patent statistics

Patent statistics: which date?

2. Application date
(date of filing)
☺ easily available
☹ depends on application strategy
(where to file first)

Recommended use:
- investigation of patenting strategies
- in other cases: only when priority date not available

Patent statistics: which date?

3. Publication of application
(earliest publication of invention)

☺ shows when invention became known to public
☹ at least 18 month after invention was made
☹ different rules in different countries

Recommended use:
- to assess influence of inventions (trend analysis etc.)
- assessment of state of the art

Patent statistics: which date?

4. Latest publication date

☺ Shows amendments to document

☹ Related to procedural issues rather than the invention

☹ Distorted statistics due to republication of documents

Patent statistics: which date?

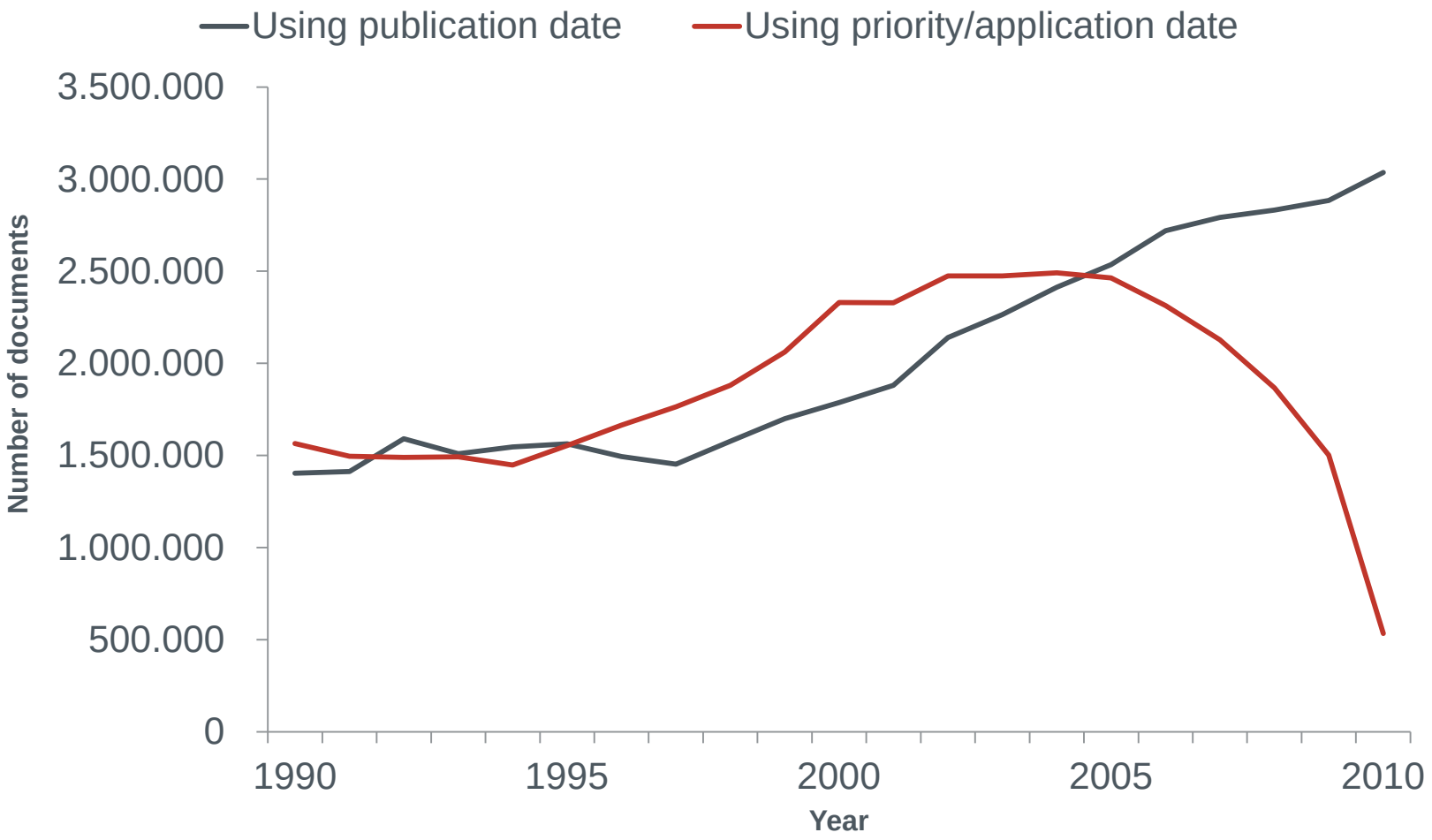
5. Patent grant (issue of patent)

☺ available for recent years

☹ depends on procedural tactics (☺)

☹ depends on performance of patent offices

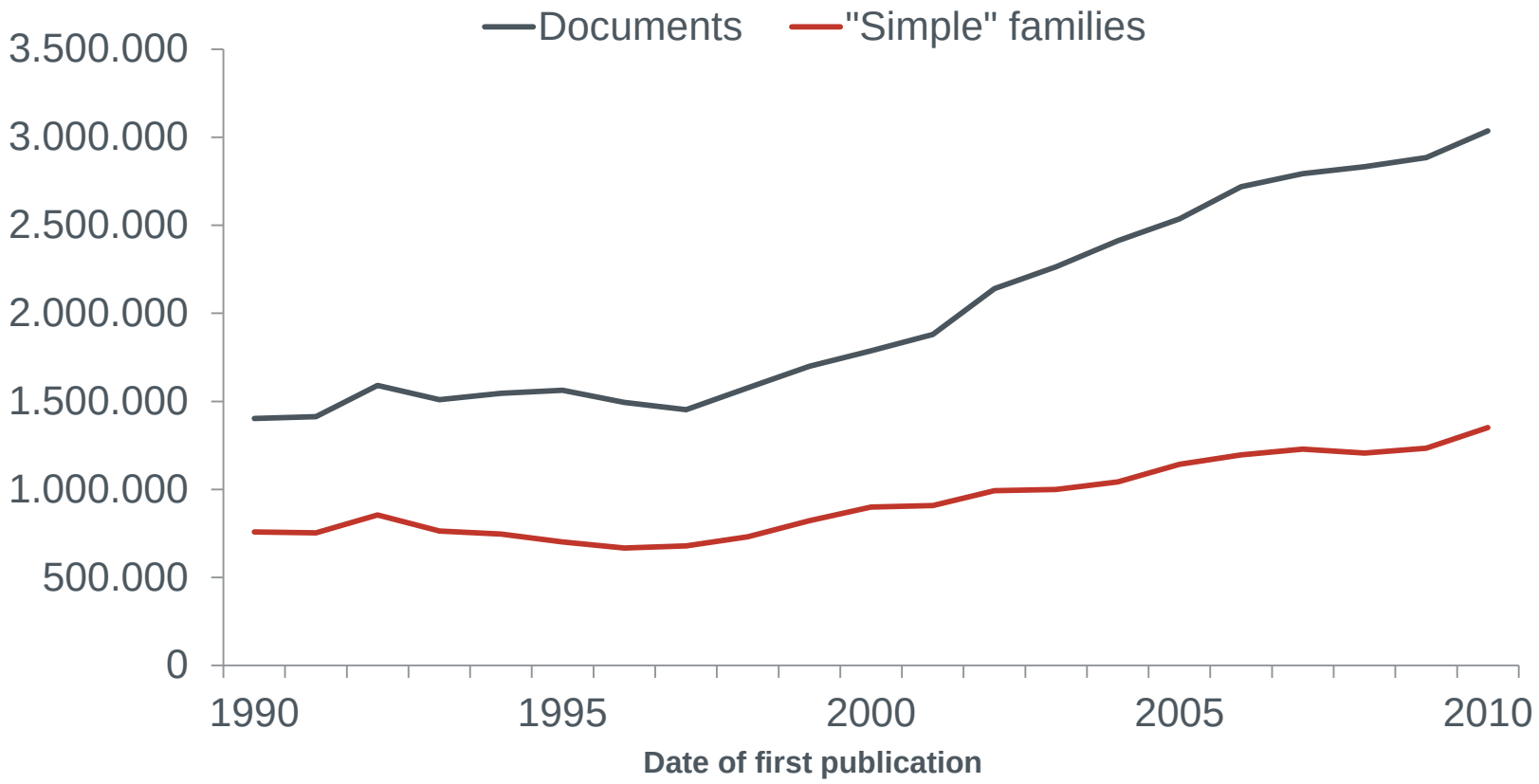
Example trend analysis: differences in priority/application date or publication date



Source: Worldwide patent database DOCDB, as of July 2011.

Counting patent families instead of documents

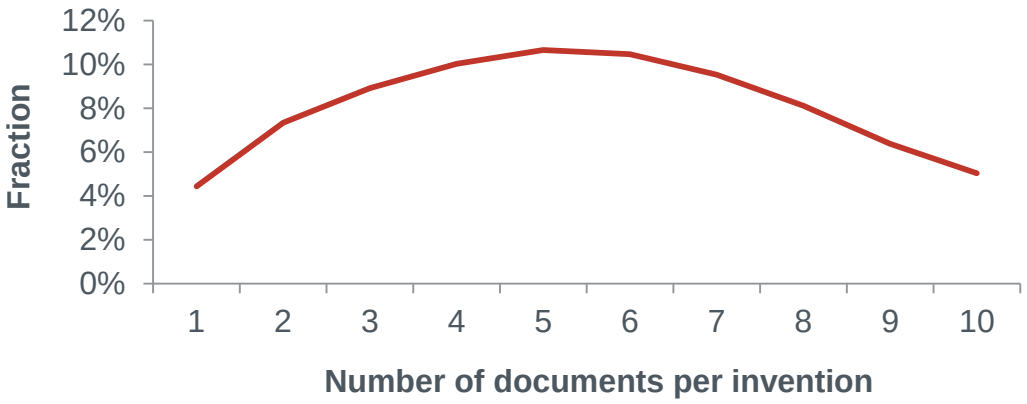
- Patent family: comprises all worldwide patent filings pertaining to the same invention



Source: DOCDB, as of July 2011.

Counting patent documents: disadvantages

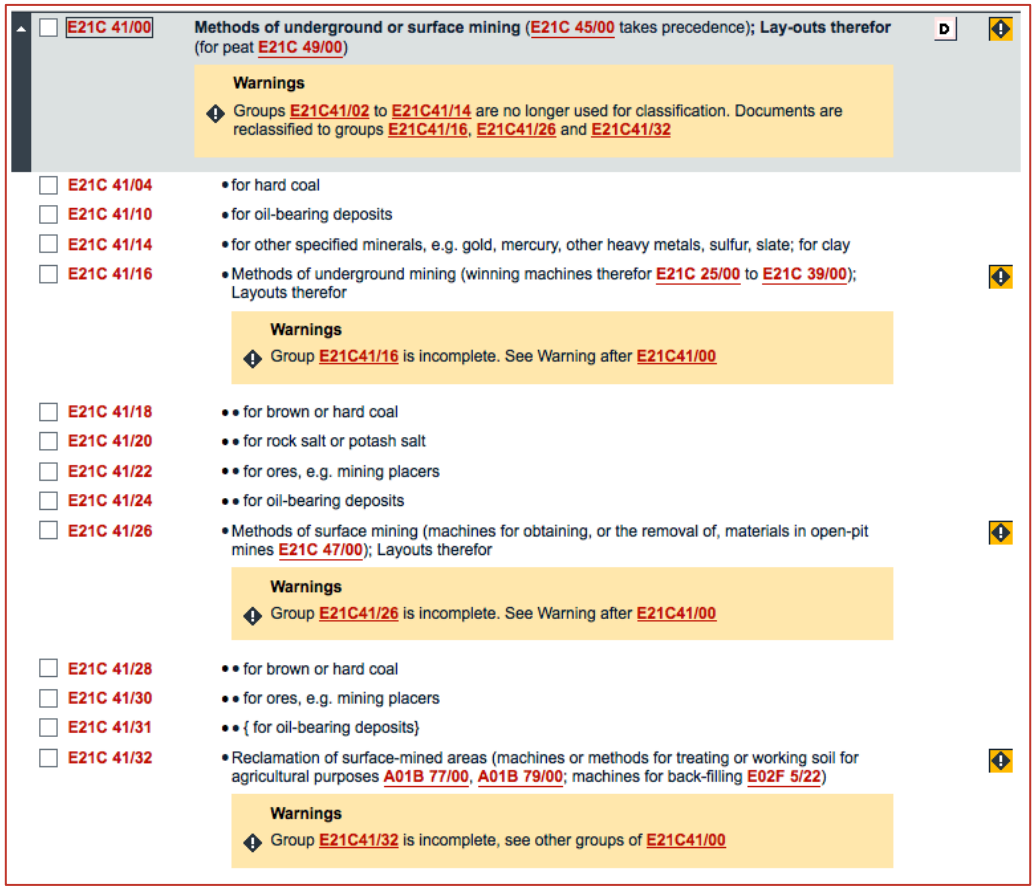
- There may be several documents per patent:
  - patent application and granted patent
  - amendments, search reports, …
- Often more than one patent application filed per invention
  - international route, regional route, national route
  - patent system-specific aspects, e.g. US provisional applications, continuation-in-parts, etc.
- Example: Inventions with at least 1 EP patent family member:



Source: DOCDB, as of July 2011. Documents per DOCDB simple family, from 1990 on-wards.

Patent statistics with patent classes: typical kinds of artefacts

- Ongoing reclassification: Example CPC classification:



E21C 41/00 **Methods of underground or surface mining** (E21C 45/00 takes precedence); **Lay-outs therefor** (for peat E21C 49/00) D

Warnings

Groups E21C41/02 to E21C41/14 are no longer used for classification. Documents are reclassified to groups E21C41/16, E21C41/26 and E21C41/32

E21C 41/04 • for hard coal

E21C 41/10 • for oil-bearing deposits

E21C 41/14 • for other specified minerals, e.g. gold, mercury, other heavy metals, sulfur, slate; for clay

E21C 41/16 • Methods of underground mining (winning machines therefor E21C 25/00 to E21C 39/00); Layouts therefor

Warnings

Group E21C41/16 is incomplete. See Warning after E21C41/00

E21C 41/18 • • for brown or hard coal

E21C 41/20 • • for rock salt or potash salt

E21C 41/22 • • for ores, e.g. mining placers

E21C 41/24 • • for oil-bearing deposits

E21C 41/26 • Methods of surface mining (machines for obtaining, or the removal of, materials in open-pit mines E21C 47/00); Layouts therefor

Warnings

Group E21C41/26 is incomplete. See Warning after E21C41/00

E21C 41/28 • • for brown or hard coal

E21C 41/30 • • for ores, e.g. mining placers

E21C 41/31 • • { for oil-bearing deposits}

E21C 41/32 • Reclamation of surface-mined areas (machines or methods for treating or working soil for agricultural purposes A01B 77/00, A01B 79/00; machines for back-filling E02F 5/22)

Warnings

Group E21C41/32 is incomplete, see other groups of E21C41/00

Fractional values

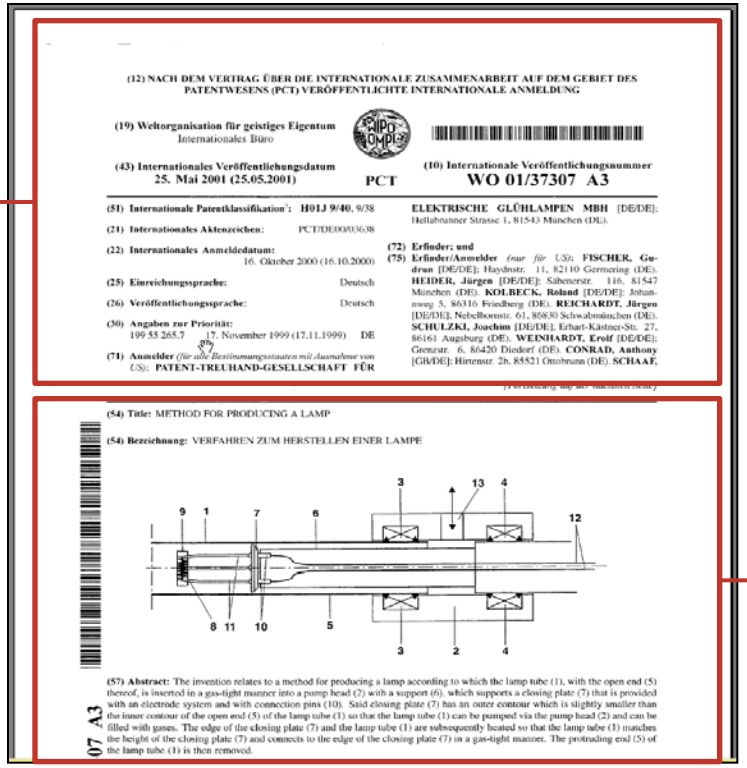
- Help to avoid multiple counting and overvaluation
- Example: CPC classes of EP2704902 A1 ('Magnetic graphene-like nanoparticles or graphitic nano-or microparticles and method of production and uses thereof')

CPC class
A61K49/10
B32B9/00
B82Y5/00
C01B31/0446
C01B31/0476
C01B2204/20
Σ

Patent statistics: procedure

Basic search → Statistical analysis → Processing results

Traditional statistical analysis of structured Information



(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
25. Mai 2001 (25.05.2001)

PCT

(10) Internationale Veröffentlichungsnummer
WO 01/37307 A3

(51) Internationale Patentklassifikation⁷: H01J 9/40, 9/38

(21) Internationales Aktenzeichen: PCT/DE00/03638

(22) Internationales Anmeldedatum:
16. Oktober 2000 (16.10.2000)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
199 55 265.7 17. November 1999 (17.11.1999) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): PATENT-TREUHAND-GESELLSCHAFT FÜR ELEKTRISCHE GLÜHLAMPEN MBH [DE/DE], Hellabrunner Strasse 1, 815-43 München (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): FISCHER, Gudrun [DE/DE]; Haydnstr. 11, 82110 Germering (DE). HEIDER, Jürgen [DE/DE]; Säbenerstr. 116, 81547 München (DE). KOLBECK, Roland [DE/DE]; Johannweg 5, 86316 Friedberg (DE). REICHARDT, Jürgen [DE/DE]; Nebelbornstr. 61, 86830 Schwabmünchen (DE). SCHULZKI, Joachim [DE/DE]; Erhart-Kästner-Str. 27, 86161 Augsburg (DE). WEINHARDT, Erolf [DE/DE]; Grenzstr. 6, 86420 Diedorf (DE). CONRAD, Anthony [GB/DE]; Hirtenstr. 2b, 85521 Ottobrunn (DE). SCHAAF,

(54) Title: METHOD FOR PRODUCING A LAMP

(54) Bezeichnung: VERFAHREN ZUM HERSTELLEN EINER LAMPE

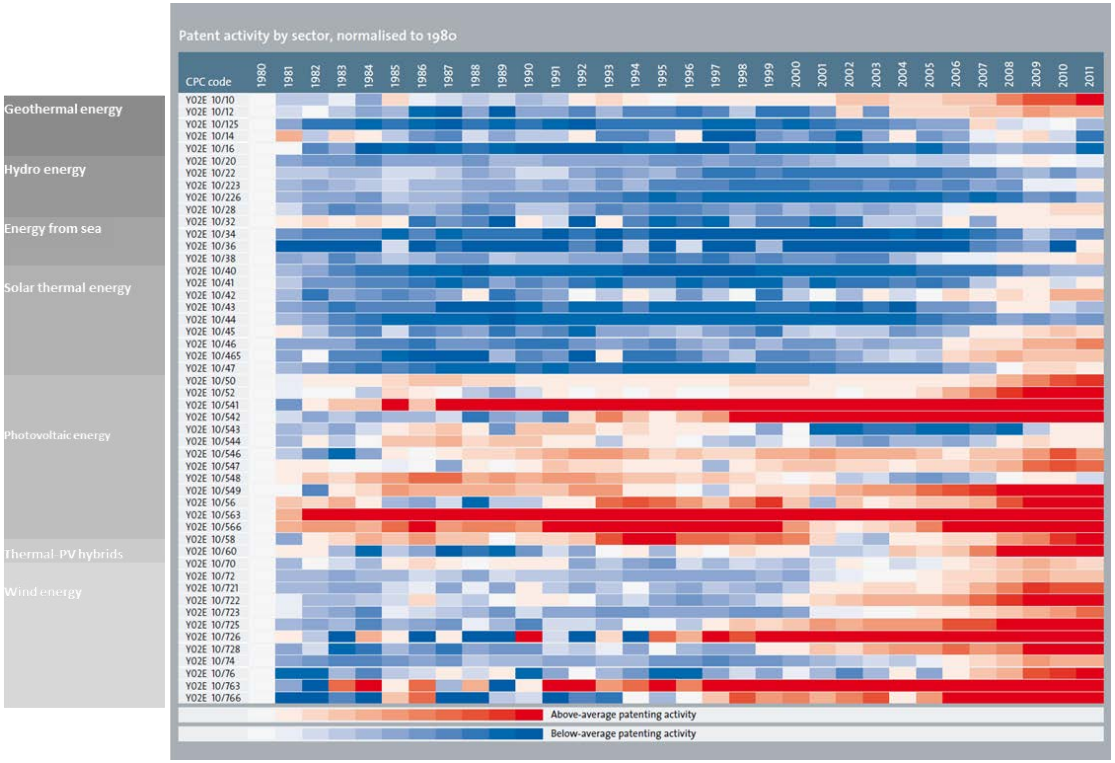
(57) Abstract: The invention relates to a method for producing a lamp according to which the lamp tube (1), with the open end (5) thereof, is inserted in a gas-tight manner into a pump head (2) with a support (6), which supports a closing plate (7) that is provided with an electrode system and with connection pins (10). Said closing plate (7) has an outer contour which is slightly smaller than the inner contour of the open end (5) of the lamp tube (1) so that the lamp tube (1) can be pumped via the pump head (2) and can be filled with gases. The edge of the closing plate (7) and the lamp tube (1) are subsequently heated so that the lamp tube (1) matches the height of the closing plate (7) and connects to the edge of the closing plate (7) in a gas-tight manner. The protruding end (5) of the lamp tube (1) is then removed.

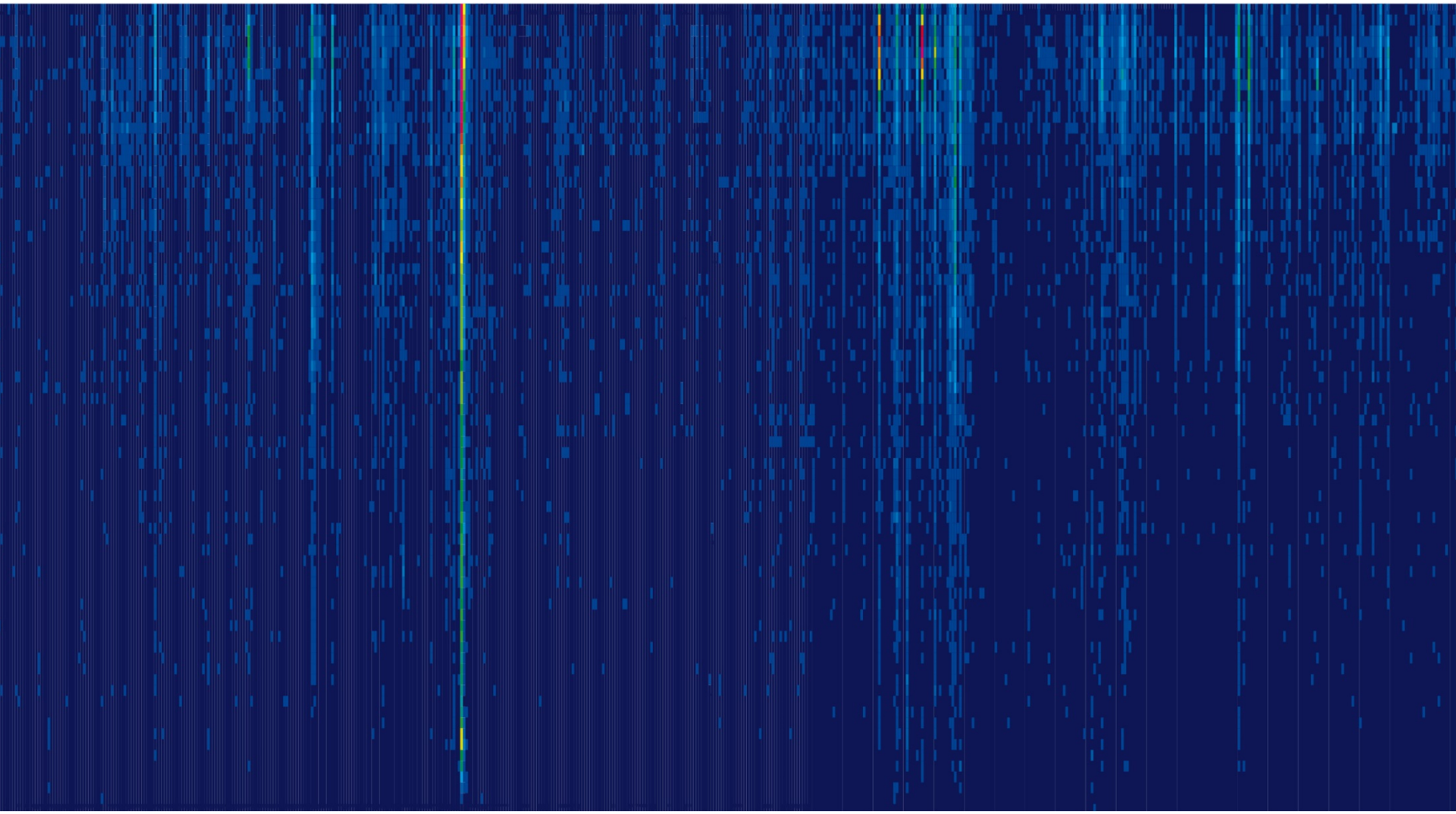
07 A3

Data mining and text mining of unstructured information

Visualisation of results

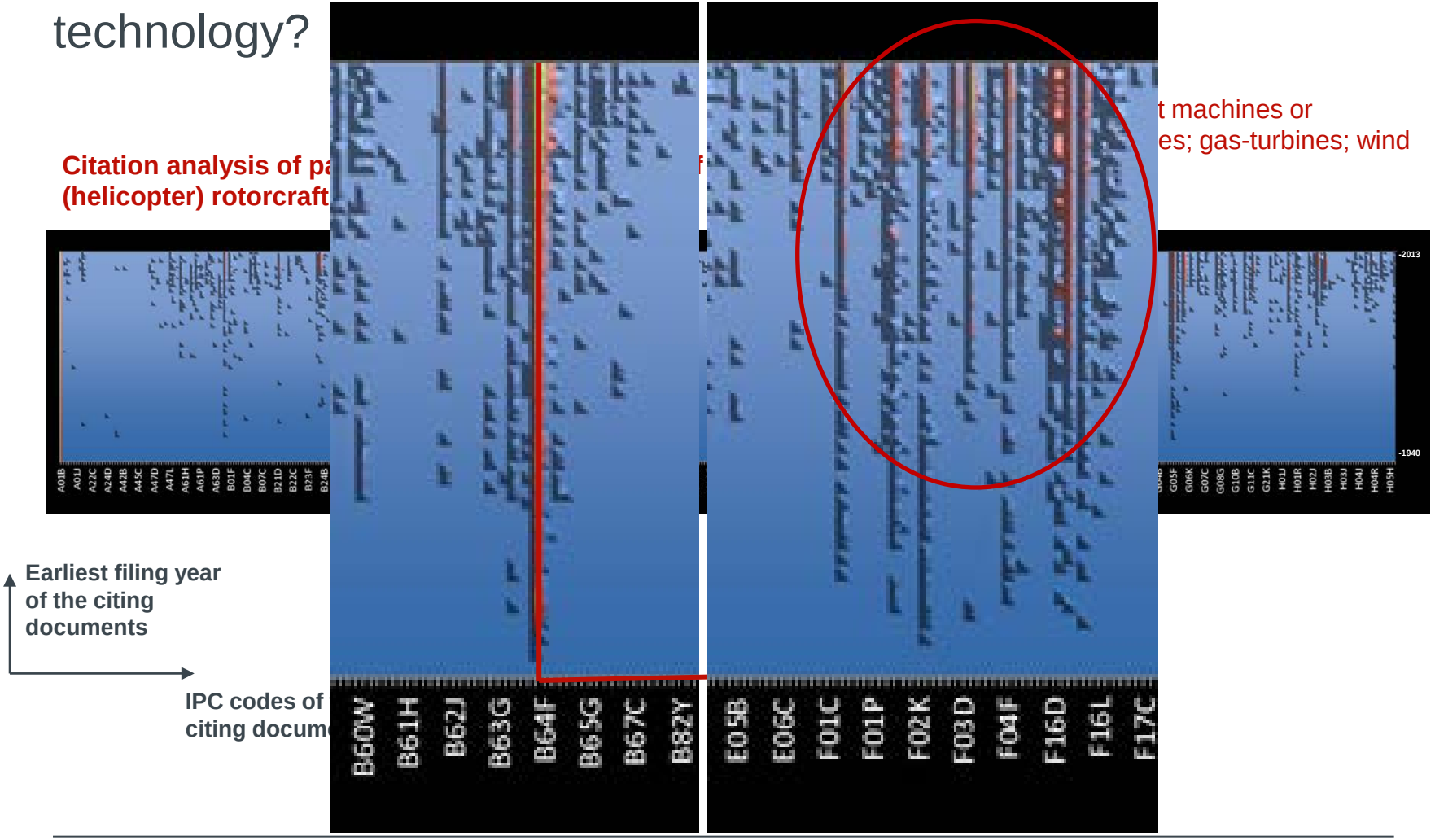
- Objective: to enable client to easily understand results
- e.g. heatmaps: Patent activity in the field of renewable energy technologies:





What patent analyses can do

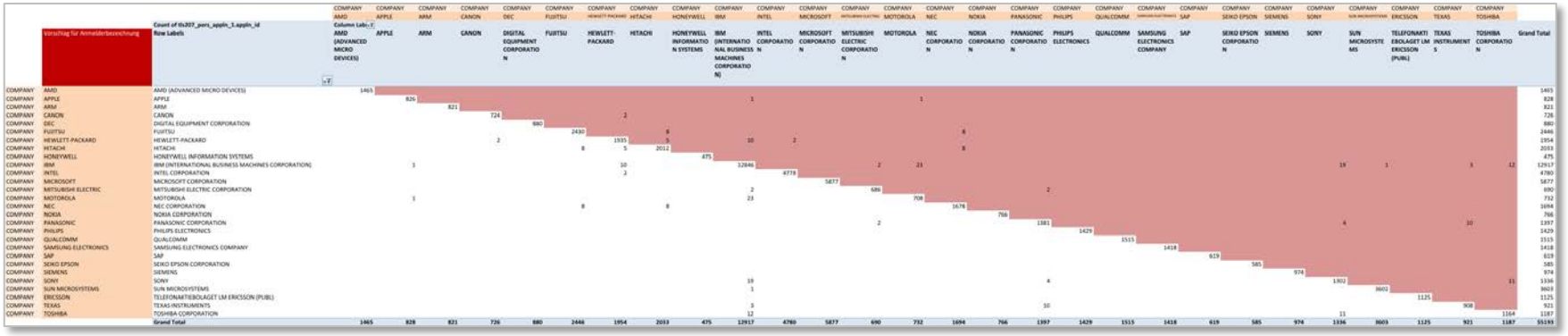
Example: In which technical fields may I also use/sell my technology?



The importance of visualisation

Which is clearer? This?

Data table

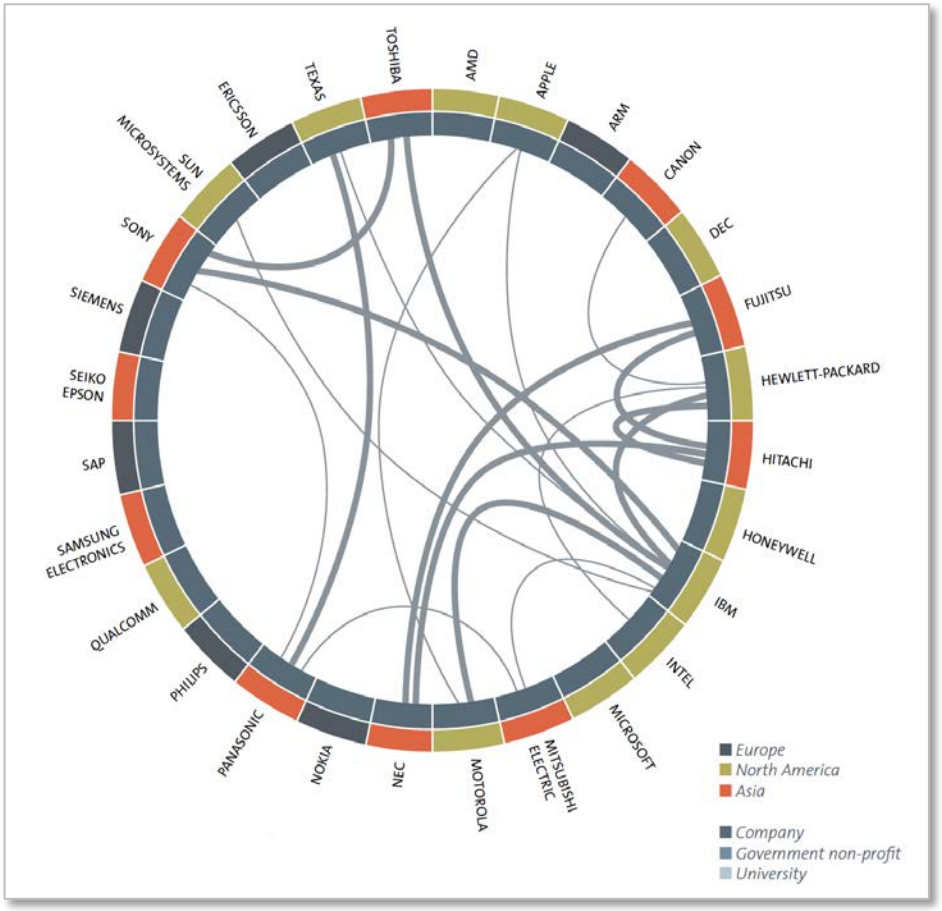


Vorschlag für Anmelder/bezeichnung	Count of tls207_pers_appln_1.appln_id	COMPANY AMD	COMPANY APPLE	COMPANY ARM	COMPANY CANON	COMPANY DEC	COMPANY FUJITSU	COMPANY HEWLETT-PACKARD	COMPANY HITACHI	COMPANY HONEYWELL	COMPANY IBM	COMPANY INTEL	COMPANY MICROSOFT	COMPANY MITSUBISHI ELECTRIC	COMPANY MOTOROLA	COMPANY NEC	COMPANY NOKIA	COMPANY PANASONIC	COMPANY PHILIPS	COMPANY QUALCOMM	COMPANY SAMSUNG ELECTRONICS COMPANY	COMPANY SAP	COMPANY SEIKO EPSON	COMPANY SIEMENS	COMPANY SONY	COMPANY SUN MICROSYSTEMS	COMPANY TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)	COMPANY TEXAS INSTRUMENTS	COMPANY TOSHIBA	Grand Total
COMPANY	AMD	1465										1				1														1465
COMPANY	APPLE		826																											828
COMPANY	ARM			821																										821
COMPANY	CANON				724				2																					726
COMPANY	DEC					880																								880
COMPANY	FUJITSU						2430		8																					2446
COMPANY	HEWLETT-PACKARD				2			1935	5			10	2				8													1954
COMPANY	HITACHI						8	5	2012								8													2033
COMPANY	HONEYWELL									475																				475
COMPANY	IBM			1				10			12846			2	23										19	1	1	12		12917
COMPANY	INTEL							2				4778																		4780
COMPANY	MICROSOFT												5877																	5877
COMPANY	MITSUBISHI ELECTRIC										2			686					2											690
COMPANY	MOTOROLA			1							23				708															732
COMPANY	NEC						8		8							1678														1694
COMPANY	NOKIA																766													766
COMPANY	PANASONIC																	1381												1397
COMPANY	PHILIPS																		1429											1429
COMPANY	QUALCOMM																			1515										1515
COMPANY	SAMSUNG ELECTRONICS																				1418									1418
COMPANY	SAP																					619								619
COMPANY	SEIKO EPSON																						585							585
COMPANY	SIEMENS																							974						974
COMPANY	SONY											19						4							1302				11	1336
COMPANY	SUN MICROSYSTEMS											1														3602				3603
COMPANY	ERICSSON																										1125			1125
COMPANY	TEXAS											3						10										908		921
COMPANY	TOSHIBA											12													11				1164	1187
	Grand Total	1465	828	821	726	880	2446	1954	2033	475	12917	4780	5877	690	732	1694	766	1397	1429	1515	1418	619	585	974	1336	3603	1125	921	1187	55193

The importance of visualisation

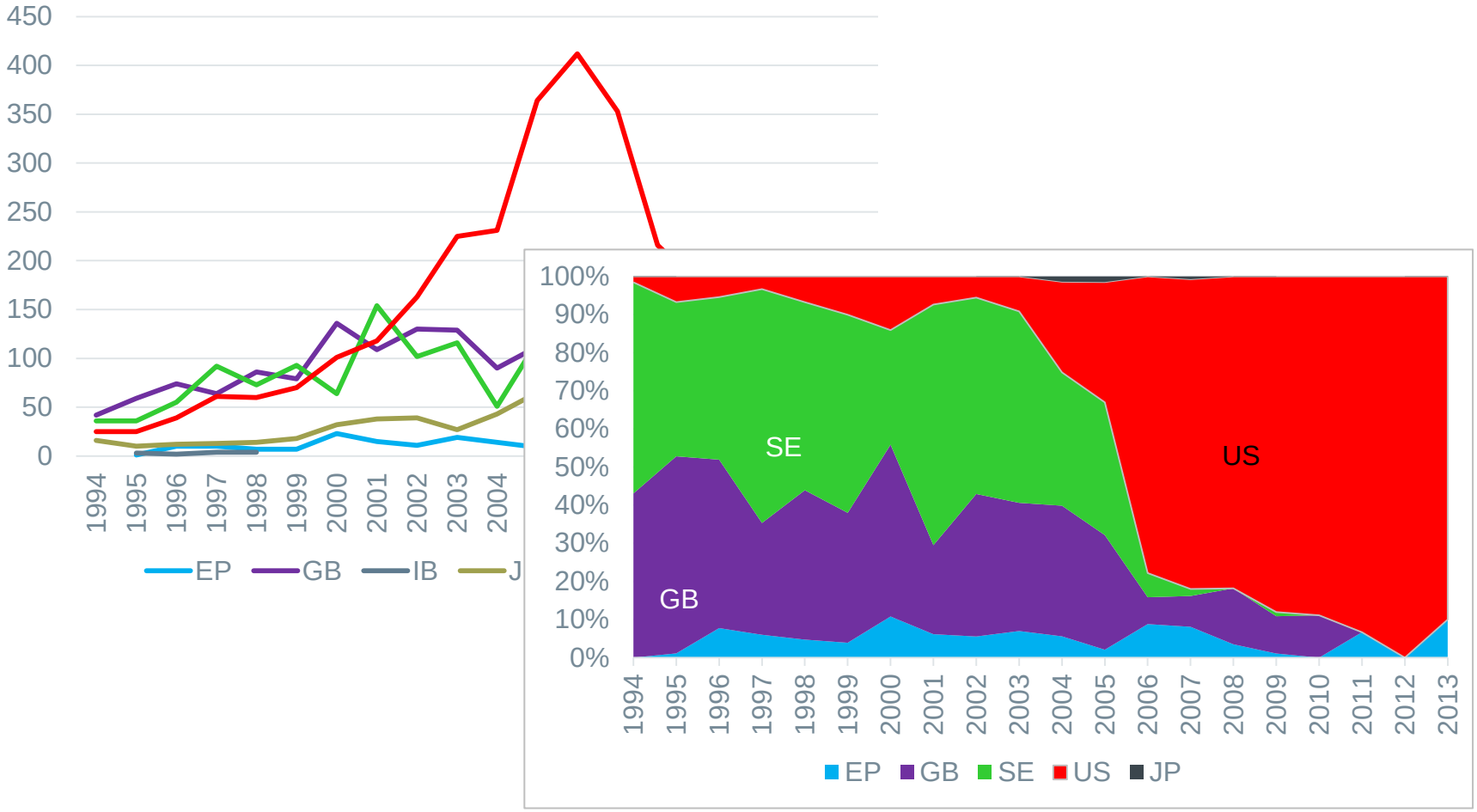
Which is clearer? Or this?

Co-applicant chart



The importance of visualisation

The choice of visualisation is important



Conclusions

- Know what you're analysing and why
- Take your decisions on which data to use in an informed way
- Don't over-analyse
- Help a decision-maker by
  - clearly describing what you've done
  - explaining your data and analysis choices
  - presenting the results in a way that supports the objective