

■ Name harmonization and classification issues

Development and effective use of IP statistics
Manila, December 17-19, 2013

Mosahid Khan
Head, IP Statistics

Overview

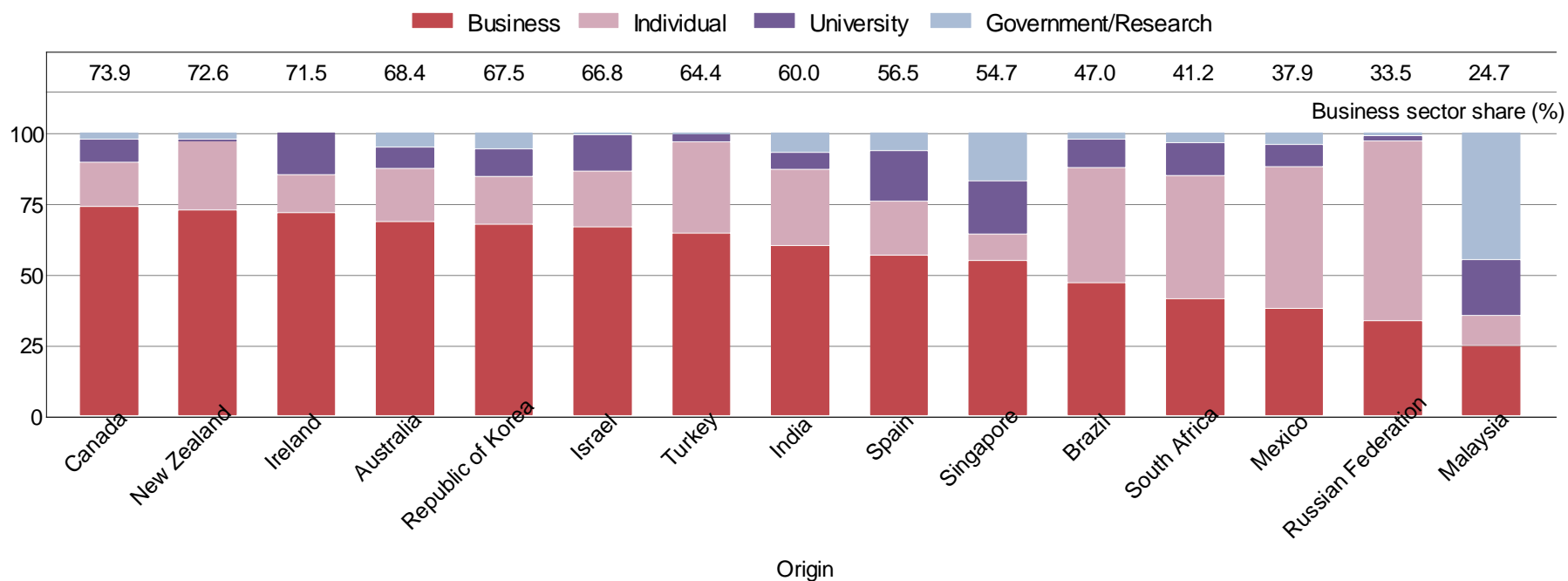
- Why do we need IP statistics by technology, institutional/industrial sectors, companies, etc.?
- How to develop such data - methodologies
- Future (methodological) work

Why do we need IP statistics broken-down by different criteria?

Institutional sectors

■ Method: algorithm/key-word search used to determine sectors

■ WIPO, OECD, Eurostat, etc.

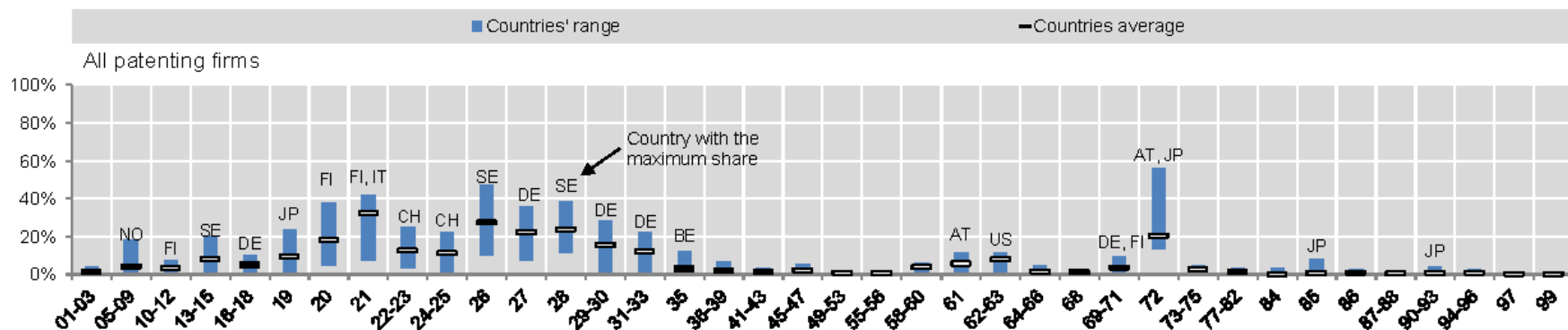


Industrial sectors - patents

■ Method: concordance table / firm level data

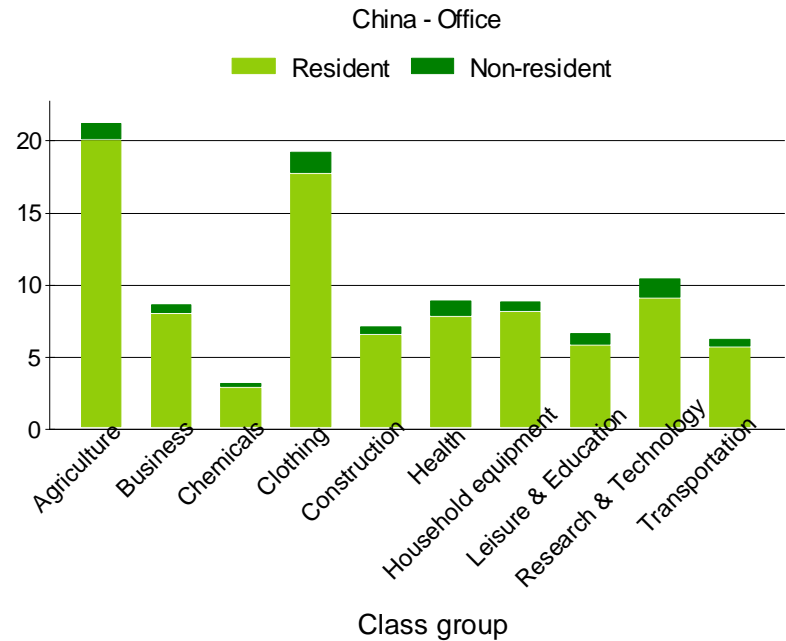
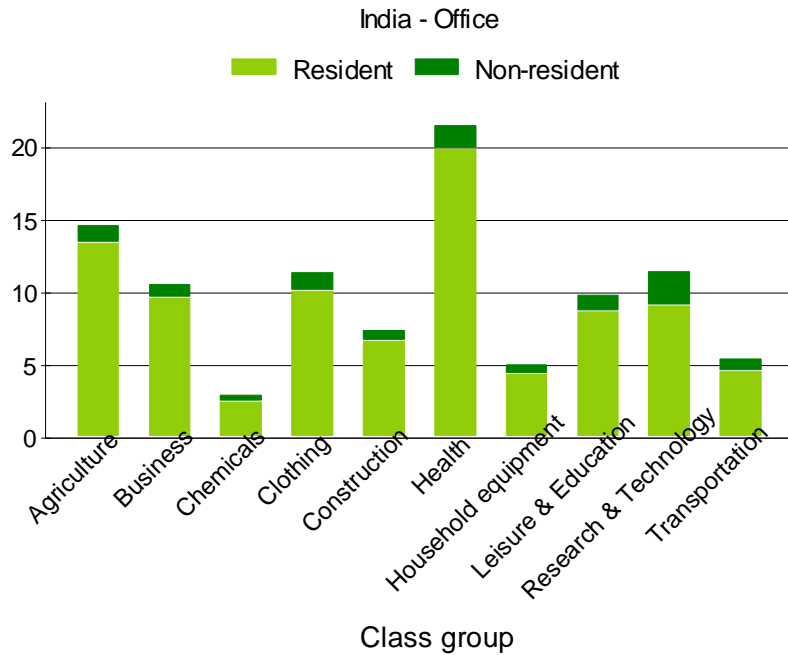
■ Yale, MERIT, USPTO, Johnson/OECD, EC.

Figure 10. Proportion of patenting firms in total, by industry (ISIC, rev.4) and size class



Squicciarini and Dernis (2013)

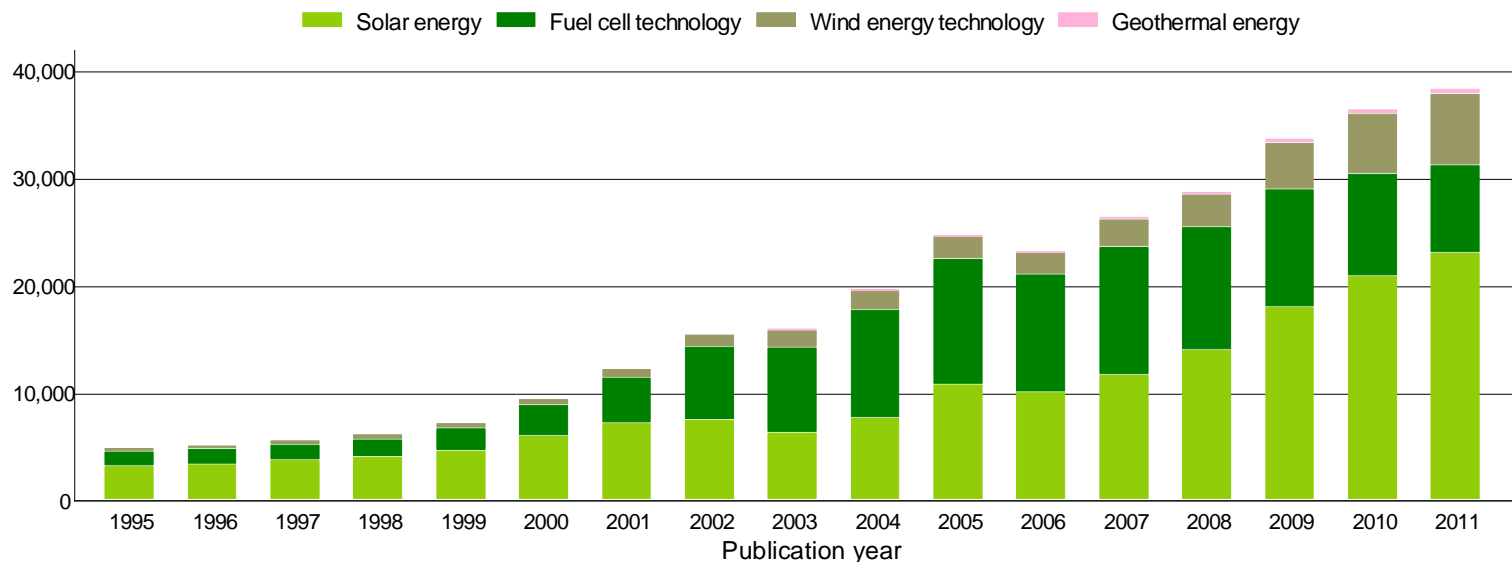
Industrial sectors - marks



Technological fields - I

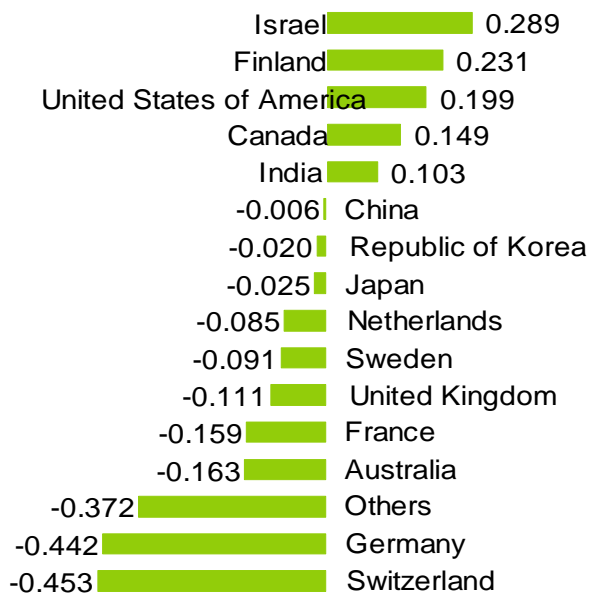
■ Method: concordance table / firm level data / Key-word search, etc.

■ WIPO IPC-Technology

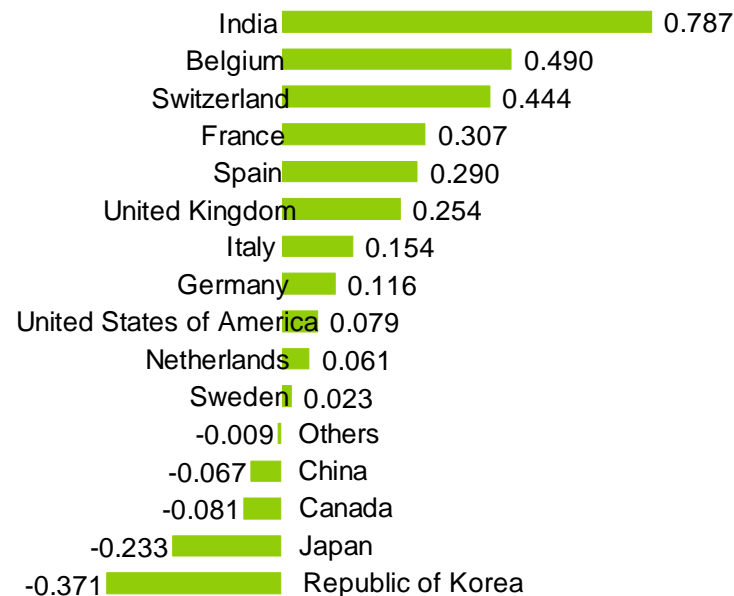


Technological fields - II

Computer technology



Organic fine chemistry



Name harmonization

■ Method: algorithm

■ WIPO, EPO, OECD, USPTO, Eurostat, etc.

■ Thoma *et al.* (NBER WP), Magerman *et al.* (2006), etc.

KONINKIJE PHILIPS ELECTRONICS N.V.	PHILIPS ELECTRONICS, N.V.
KONINKLIJKE PHILIPS ELECTRONICS N.V	KONINKLIJKE PHILIPS ELECTRONICS, N. V.
KONINKLIJKE PHILIPS ELECTRONICS NV	KONINKLIJEKE PHILIPS ELECTRONICS N.V.
KONINLIJKE PHILIPS ELECTRONICS N.V.	KONINKLIJKE PHILIPS ELECTZRONICS N.V.
KONINKLIJKE PHILIPS ELECTRONICS, N.V.	KONINKLIJKE PHILIPPS ELECTRONICS N.V.
KONINKLIJKE PHILIPS ELECTRONICS N. V.	KONINKLIJKJE PHILIPS ELECTRONICS N.V.
KONINKLIJKE PHILIPS ELECTRONICS N.V.,	KONINKLIJKE PHILLIPS ELECTRONICS N.V.
KONINKLIKE PHILIPS ELECTRONICS N.V.	KONINKLIJKE PHLIPS ELECTRONICS N.V.
KONINKLIJKE PHILIPS ELECRONICS N.V.	KONINGLIJKE PHILIPS ELECTRONICS N.V.
KONINKLIJKE PHILIPS ELECTRONIQUE N.V.	KONINKLIJKE PHILIPS ELECTRONIC NV
KONINKLIJKE PHILIPS ELECTRONICS	PHILIPS ELECTRONIC N.V.
KONINKLIGKE PHILIPS ELECTRONICS N.V.	KONINKLIJKE PHILIPS ELECTRONIC N. V.
KONINKLIJKC PHILIPS ELECTRONICS N.V.	KRONINKLIJKE PHILIPS ELECTRONICS
KONINKLIJKE PHILIPS ELECTRONICS NV	KONINKLIJKE PHILIPS ELECTRPMOCS. N.V.
NINKLIJKE PHILIPS ELECTRONICS NV	KONINKLIJKE PHILIPS ELECTRONICS B.V.
KNONINKLIJKE PHILIPS ELECTRONICS, N. V.	KONINKLIJKE, PHILIPS
KONINKLIJKE PHILIPS ELECTRONICS N V	KONLNKLIJKE PHILIPS ELECTRONICS N.V.
KONINKLIJKE PHILIPS ELECTRONICS, NV	KONINKLIJKE PHILIPS ELECTTRONICS N.V.
PHILIPS ELECTRONICS N.V.	KONINKLIJKE PHILIPS ELECTRONICS N.V.
PHILIPS ELECTRONICS N.V	KONINKLIJKE PHILIPS N.V.

Name harmonization – WIPO method

PCT name cleaning:

1. Formatting:

- a. Capitalizing.
- b. Removing special characters, such as “.”, “-” etc.
- c. Removing common words, such as “the”, “of”, “and” etc.
- d. Standardizing company type, such as “ltd” to “limited” etc.
- e. Reversing some names, such as “university of xxx” to “xxx university”.
- f. Calculating similarity between all name pairs (Levenshtein distance)
- g. Calculating similarity between all address pairs (Levenshtein distance)

2. Matching:

- a. Automatic matching by standard names.
- b. Automatic matching by matched name variations.
- c. Semi-automatic (user confirmation) matching by maximum similarity of names.
- d. Semi-automatic matching by maximum similarity of addresses.
- e. Semi-automatic matching by key words assigned to top applicants.

3. Issues:

- a. Do not match names with different company type.
- b. Check address before matching names with different entity names.
- c. Do not set up similarity threshold and apply automatic matching according to similarity.

Name harmonization – WIPO example

Ranking	Applicant's name	Origin	PCT 2012
1	ZTE CORPORATION	China	3,906
2	PANASONIC CORPORATION	Japan	2,951
3	SHARP KABUSHIKI KAISHA	Japan	2,001
4	HUAWEI TECHNOLOGIES CO., LTD.	China	1,801
5	ROBERT BOSCH CORPORATION	Germany	1,775
6	TOYOTA JIDOSHA KABUSHIKI KAISHA	Japan	1,652
7	QUALCOMM INCORPORATED	United States of America	1,305
8	SIEMENS AKTIENGESELLSCHAFT	Germany	1,272
9	KONINKLIJKE PHILIPS ELECTRONICS N.V.	Netherlands	1,230
10	TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)	Sweden	1,197
11	LG ELECTRONICS INC.	Republic of Korea	1,094
12	MITSUBISHI ELECTRIC CORPORATION	Japan	1,042
13	NEC CORPORATION	Japan	999
14	FUJIFILM CORPORATION	Japan	891
15	HITACHI, LTD.	Japan	745
16	SAMSUNG ELECTRONICS CO., LTD.	Republic of Korea	683
17	FUJITSU LIMITED	Japan	671
18	NOKIA CORPORATION	Finland	670
19	BASF SE	Germany	644
20	INTEL CORPORATION	United States of America	640
21	HEWLETT-PACKARD DEVELOPMENT COMPANY, L.	United States of America	620
22	3M INNOVATIVE PROPERTIES COMPANY	United States of America	586
23	SONY CORPORATION	Japan	578
24	MITSUBISHI HEAVY INDUSTRIES, LTD.	Japan	566
25	SUMITOMO CHEMICAL COMPANY, LIMITED	Japan	558
26	SANYO ELECTRIC CO., LTD.	Japan	537
27	MICROSOFT CORPORATION	United States of America	531
28	INTERNATIONAL BUSINESS MACHINES CORPORAT	United States of America	528
29	CANON KABUSHIKI KAISHA	Japan	480
30	MURATA MANUFACTURING CO., LTD.	Japan	462

Future methodological work

- Most of the work has focused on patent
- Need to further improve methodology for IP data by industry
- Name consolidation is needed to improve IP data by companies
- Construction of patent portfolio of companies

Thank you!

Mosahid.khan@wipo.int