General Methodological Issues Development of IP Statistics

Kazuyuki Motohashi

Department of Technology Management for Innovation, University of Tokyo and RIETI
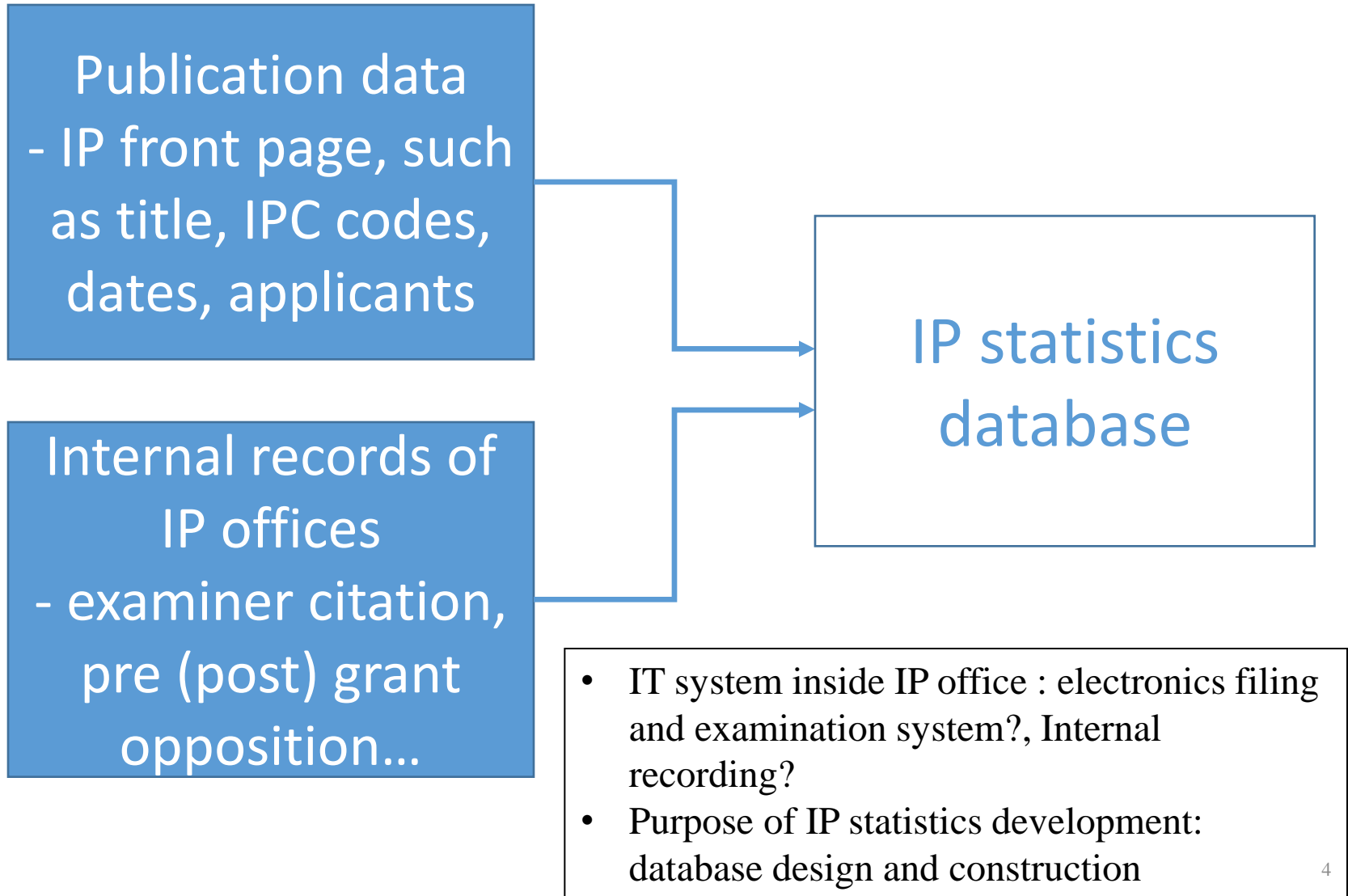
http://www.mo.t.u-tokyo.ac.jp/

Outline

1. Introduction
2. Data sources and final products: Issues to be considered
3. Example: IIP-Patent Database in Japan
4. Use of external complementary data such as PATSTAT
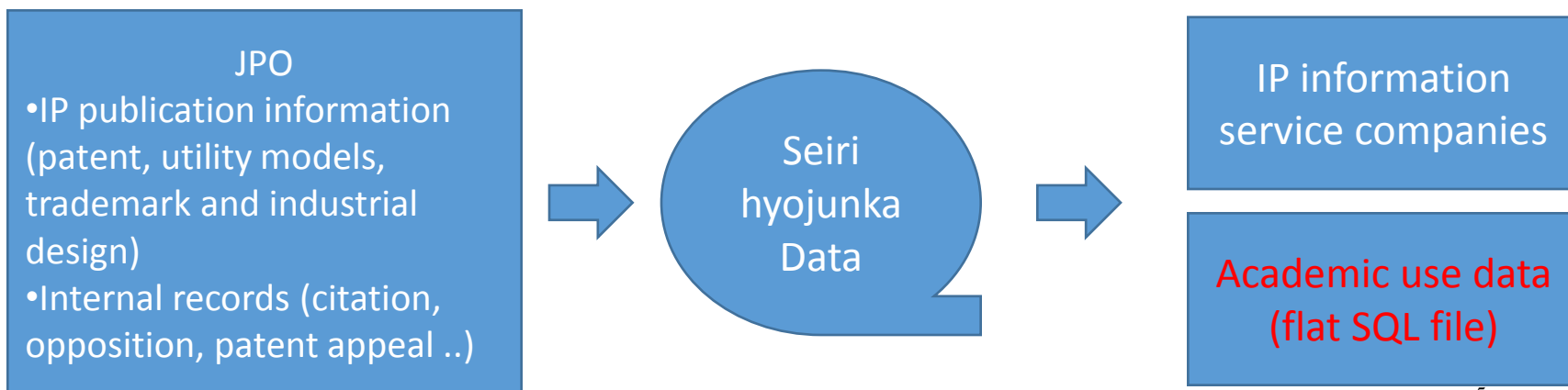
1.Introduction

- Difficulty in measuring knowledge despite the talk of "knowledge society"

- Patent (and other IP) as a source of information on technological knowledge

- Private (academic) contribution to IP based database (such as NBER patent database, EPIP, IIP …). "Open Innovation Model" for patent offices

2. Data sources and final products

Publication data
- IP front page, such as title, IPC codes, dates, applicants

Internal records of IP offices
- examiner citation, pre (post) grant opposition…

IP statistics database

- IT system inside IP office : electronics filing and examination system?, Internal recording?
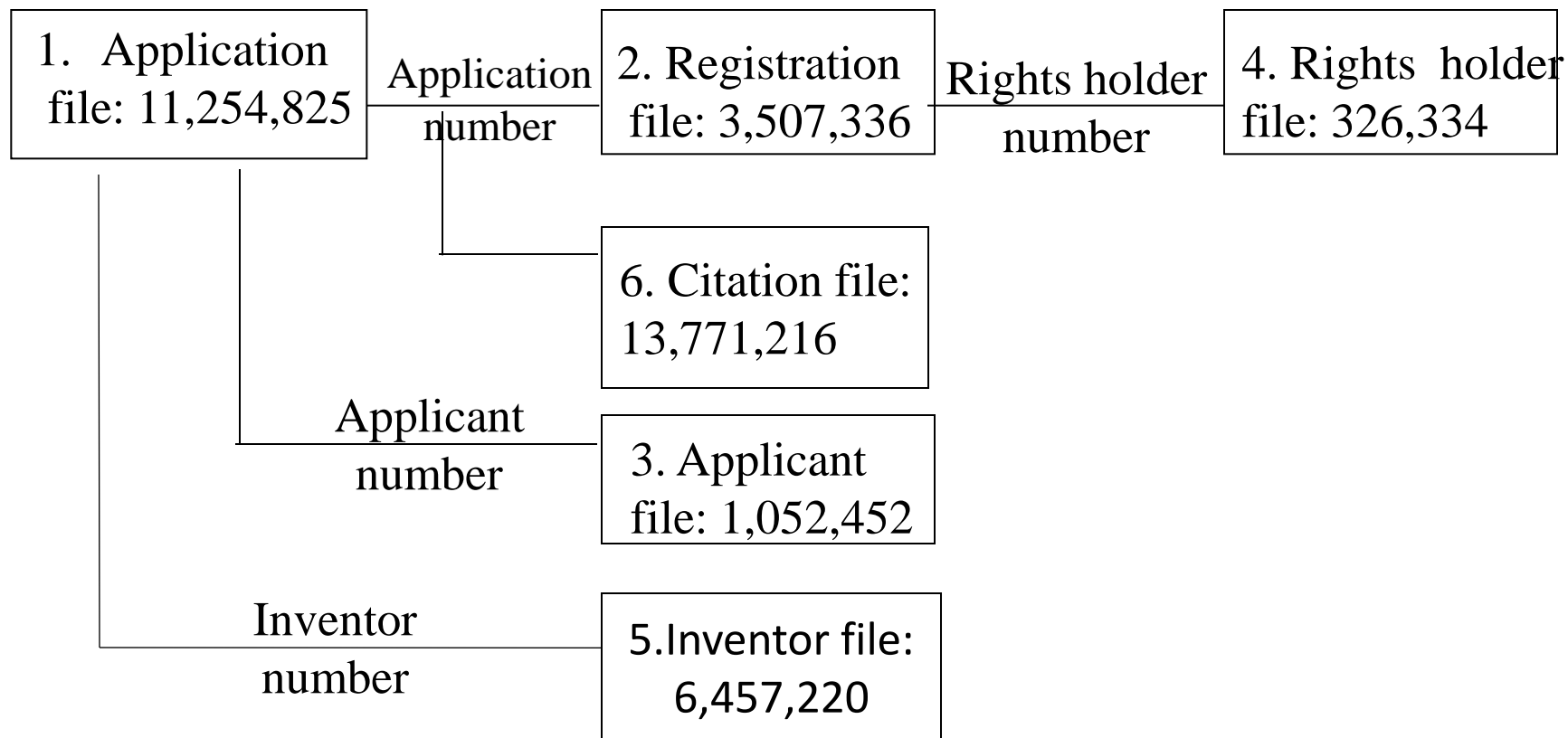- Purpose of IP statistics development: database design and construction

3. Example: IIP patent data

- Original Data Source; JPO Seiri hyojunka Data

  (Literally translating, "Arranged and Standardized Data", hereafter, JPO Patent Database)

  - It contains the information generated through acceptance of application to examination process by JPO, released 25 times a year (every two weeks).
  - About 50 boxes of half inch tape+DVD-ROM recently.

- SQL database: Voluntary initiative by academic group with a help of IIP (Institute of Intellectual Property: JPO related non profit organization) for constructing database for academic use

JPO
•IP publication information (patent, utility models, trademark and industrial design)
•Internal records (citation, opposition, patent appeal ..)

Seiri hyojunka Data

IP information service companies

Academic use data (flat SQL file)

Same as NBER, EPIP initiative

Structure of IIP Patent Database

1. Application file: 11,254,825	Application number	2. Registration file: 3,507,336	Rights holder number	4. Rights holder file: 326,334

6. Citation file: 13,771,216

Applicant number — 3. Applicant file: 1,052,452

Inventor number — 5.Inventor file: 6,457,220

• The data from 1964 patent application are included
• Original information is preserved

Importance of normalization for relational database

Patent #	Technology class	Applicant name	Applicant address
1	A61K	XXX co. ltd	Tokyo
1	C12N	XXX co. ltd	Tokyo
1	A61K	ZZZ university	Osaka
1	C12N	ZZZ university	Osaka
1	A61K	Mr. YYY	Sapporo
1	C12N	Mr. YYY	Sapppro

Patent #	IPC
1	A61K
1	C12N

Patent #	Applicant #
1	1
1	2
1	3

Applicant #	Applicant name	Applicant address
1	XXX co. ltd	Tokyo
2	ZZZ university	Osaka
3	Mr. YYY	Sapporo

①Patent application files

- Application number
- Application date
- Examination request date
- Applicant number
- Number of claims at patent application
- Lead IPC code at patent application
- Aggregated technology category

Aggregated Technology Category

Tech No	Title	Corresponding IPC			NBER
1	Agriculture	A01		(ExceptA01N)	6
2	Food Stuffs	A21	~ A24		6
3	Personal and Domestic Articles	A41	~ A47		6
4	Health and Amusement	A61	~ A63	(Except A61K)	3
5	Drugs	A61K			3
6	Separating, Mixing	B01	~ B09		1
7	Machine tools, Metalworking	B21	~ B23		5
8	Casting, Grinding, Layered Product	B24	~ B32	(Except B31)	5
9	Printing	B41	~ B44		6
10	Transporting	B60	~ B64		5
11	Packing, Lifting	B65	~ B68		5
12	Non organic chemistry, Fertilizer	C01	~ C05		1
13	Organic chemistry, Pesticides	C07	、 A01N		1
14	Organic molecule compounds	C08			1
15	Dyes, Petroleum	C09	~ C11		1
16	Biotechnology, Beer, Fermentation	C12	~ C14		3
17	Genetic Engineering	C12N15/			3
18	Metallurgy, Coating metals	C21	~ C30		5
19	Textile	D01	~ D07		6
20	Paper	D21	、 B31		6
21	Construction	E01	~ E06		6
22	Mining, Drilling	E21			6
23	Engine, Pump	F01	~ F04	、 F15	5
24	Engineering elements	F16	~ F17		5
25	Lighting, Steam generation, Heating	F21	~ F28		6
26	Weapons, Blasting	F41	~ F42	、 C06	6
27	Measurement, Optics, Photography	G01	~ G03		4
28	Clock, Controlling, Computer	G04	~ G08		2
29	Display, Information Storage, Instruments	G09	~ G12		2
30	Nuclear physics	G21			4
31	Electronics components, semiconductor	H01	~ H02	、 H05	4
32	Electronics circuit, communication tech.	H03	~ H04		2
33	Others	B81	、 B82		6

②Patent registration file

- Application number
- Registration number
- Registration date
- Rights expiration date
- Rights holder number
- Number of claims at patent registration
- Lead IPC code at patent registration
- Aggregated technology category

of observation by year

Legend: ■ application ■ registration
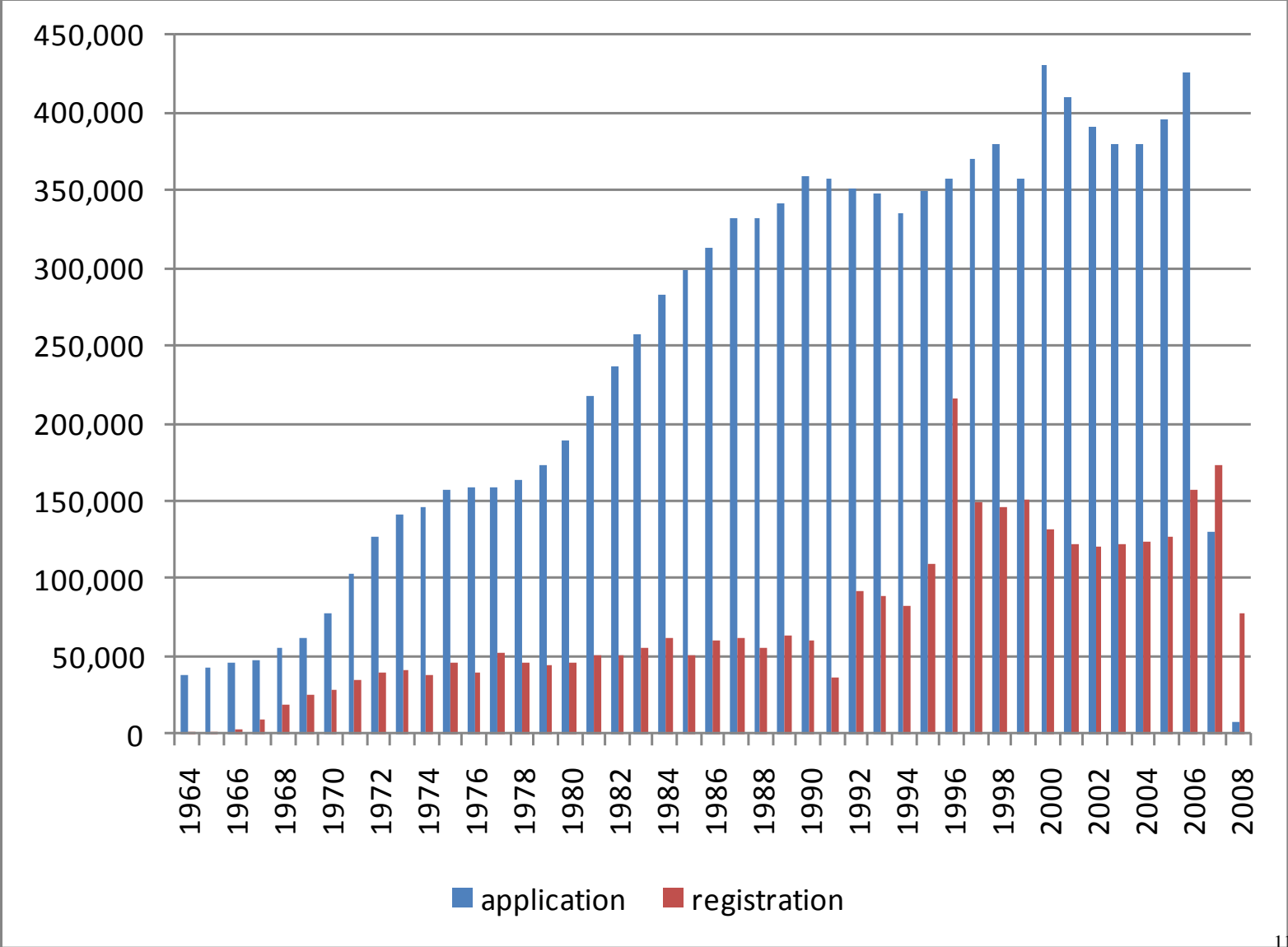
③Applicant file

- Applicant number
- Applicant name
- Applicant type (individual, corporation or government)
- Country and prefecture code
- JPO applicant code

④ Rights holder file

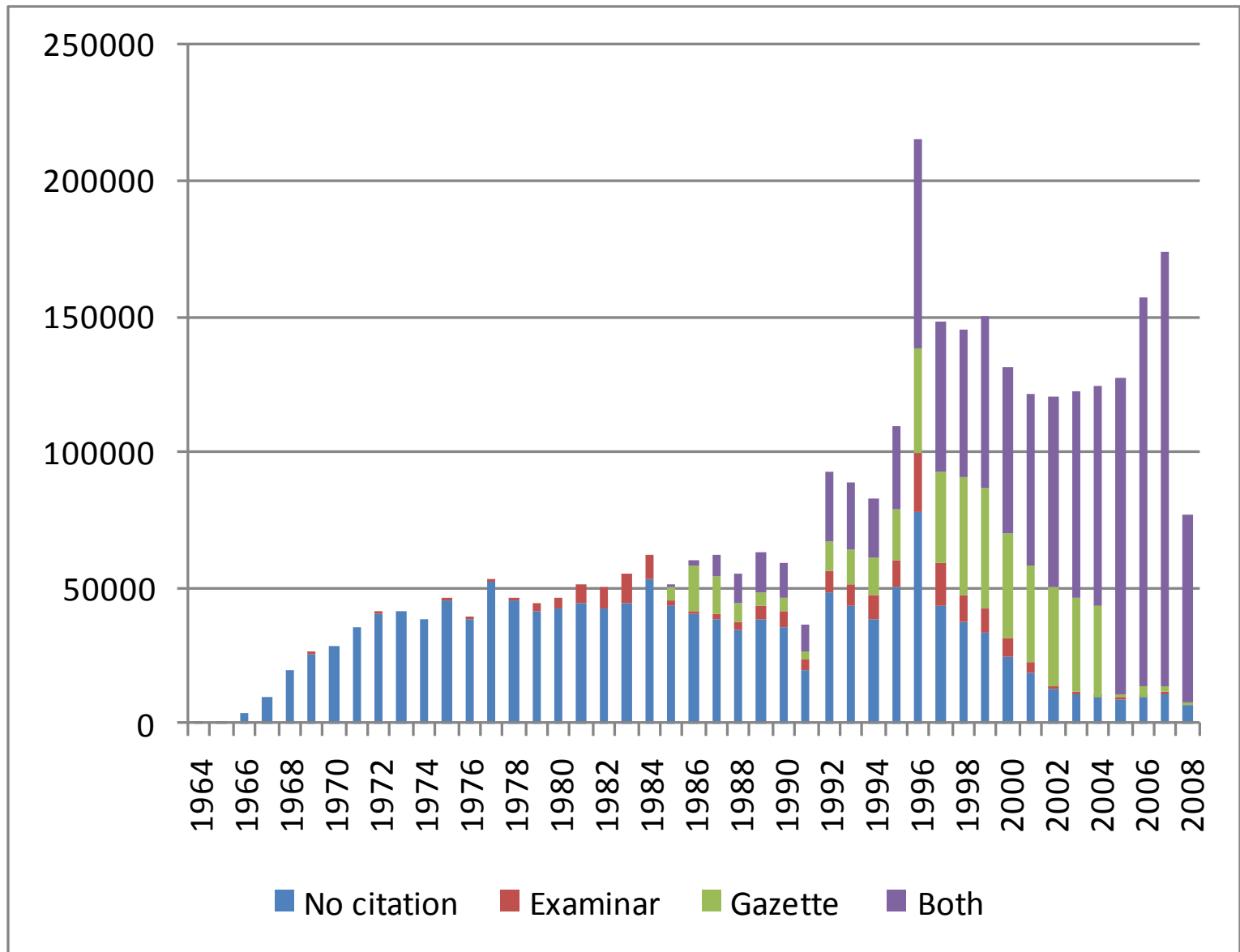- Rights holder number
- Rights holder name
- Rights holder address

⑤ Inventor file

- Inventor number
- Inventor name
- Inventor address

⑥ Citation information file

- Citing patent application number

- Cited patent application number

- Citation type

Registration patent by citation type

Legend: No citation, Examinar, Gazette, Both

4. Use of external complementary data such as PATSTAT

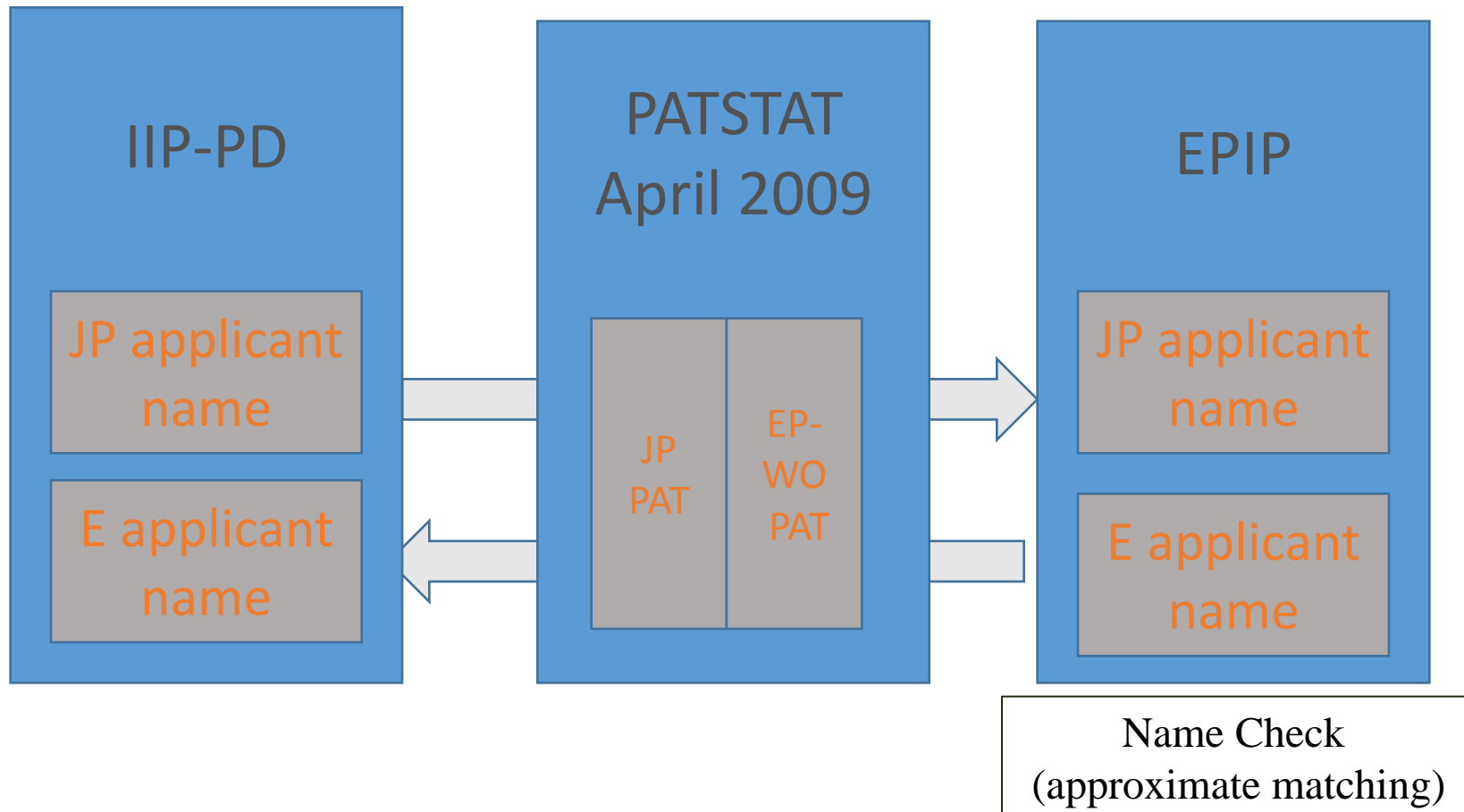Linking with patent database in other countries/regions by patent family information

◆ complementing information for internal datasets

◆ add new variables as a patent quality indicator

Useful data source : PATSTAT

◆ 80 million patent application information from over 70 patent authority

◆ Originated from DOC-DB (EPO's master documentation database)

Applicant name harmonization works for IIP-PD and EPIP database

Grid Thoma, Kazuyuki Motohashi, and Jun Suzuki, "Consolidating firm portfolios of patents across different offices. A comparison of sectoral distribution of patenting activities in Europe and Japan", IAM Discussion Paper Series #019, 2010/11

Useful sources

IIP Patent Database

http://www.iip.or.jp/e/e_patentdb/index.html

IIP PD methodological paper

Akira Goto and Kazuyuki Motohashi, "Construction of a Japanese Patent Database and a first look at Japanese patenting activities", Research Policy 36(9), 2007/11, pp.1431-1442

Patent database development works in other countries (for Asian countries, not covered well by PATSTAT)

China: Kazuyuki Motohashi, "Assessment of technological capability in science industry linkage in China by patent database", World Patent Information 30(3), September 2008, Pages 225-232

India : work undergoing

For other nations