

## Science linkages in technologies in Japan

Schumpeter Tamada, Yusuke Naito,  
Kiminori Gemba, Fumio Kodama,  
Jun Suzuki and Akira Goto

### Motivation of our study

- How is the relationship between science and technology? Is it different from technology to technology?
- How is the impact on innovation of government's support for scientific activities?



## Purpose of our study (1)

- That science serves as the engine of technological change and economic growth is widely recognized.
- Mansfield pointed out that the development of 10% of new products and process could not have been developed without substantial delay in the absence of recent academic research (Mansfield, 1991).
- As more attention has been focused on science as the source of technological change that would lead to economic growth, there has also been increased interest in the linkage between science and technological change (Narin *et al.*, 1997).

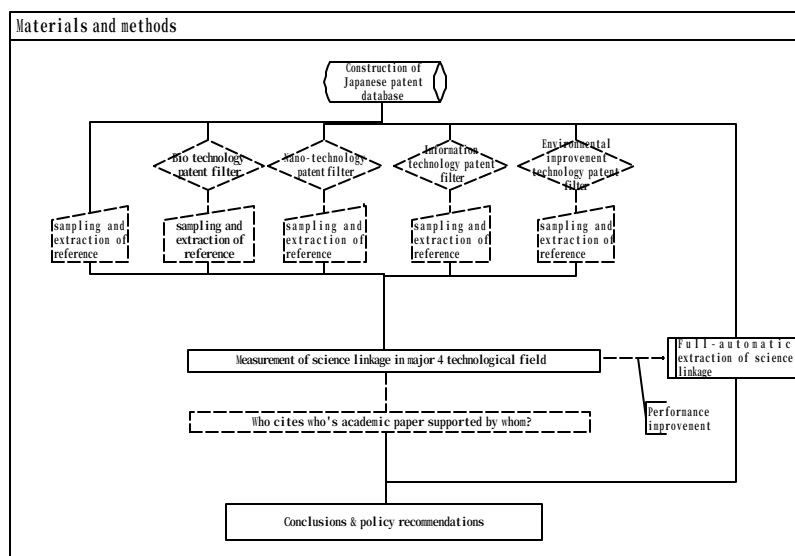


## Purpose of our study (2)

- In recent years, the “Science Linkage,” which indicates the number of non patent reference, that are mainly academic papers, cited per patent, has been used as an indicator for understanding the effect that science exerts on technology
- But no study of science linkage in Japanese patent was found

## Research questions

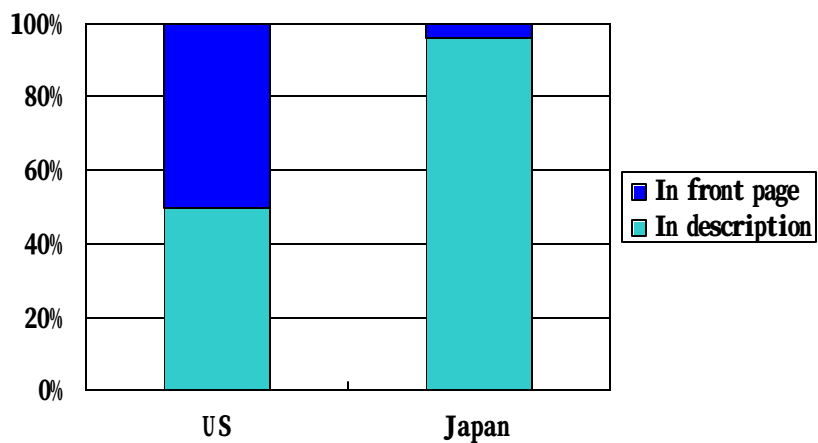
- Do inventors cite academic papers in Japanese patents? If so, is the science linkage differ from technology to technology?
- Who cites who's academic paper supported by whom?
- Is the distribution pattern of science linkage among various technological classification in Japanese patent different from that in the EU?



## Constructing original database

- Publication of unexamined applications
  - 2,380,000
  - (1993.1~2001.10)
- Patent Gazette
  - 880,000
  - (1994.1~2001.10)
- 1,100 CD-ROMs
- 800 GB

## Most references in Japanese patents are in description





## 1. Extracting cited papers and patents by hand

---



### Data used

---

- Patents granted since 1995 to 1999
- Using filtering program, patents in four priority technological fields in the Japanese Science and Technology Basic Plan was filtered out.
- Out of those patents in each technological field, 300 patents are randomly sampled.
- As control, 300 patents are randomly sampled.

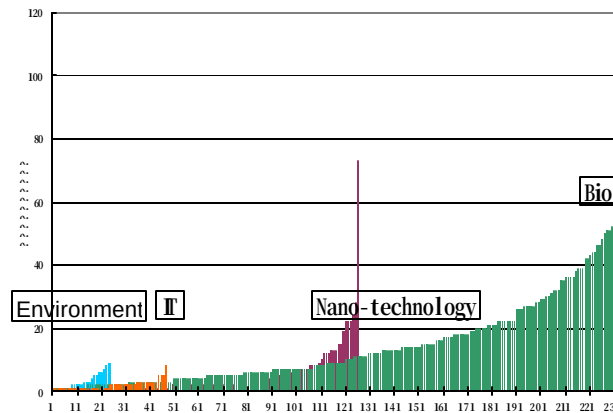
# Filtering algorithm

| Technological field     | Filtering algorithm                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Number of patents filtered |
|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|
| Bio                     | 1) IPC: C12N15+ C12N/1+ C12N/5+ C12N/7+ A61K/48<br>2) ??????????: ??????+????+????+????????+????????+?????+?????+?<br>????+????????+????+????+????????+????+?????<br>3) 1+2                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 6,965                      |
| Nano                    | 1) IPC(+FI): B82B1/00+ B82B3/00<br>2) ??????: ??+????+??????+(??*???) +????+?????+(??*??????)+(??*??<br>???) +??+????+?????+????+????+LB?+(??????*???????) +(<br>langmuir*blodgett) +????+(?????)<br>3) ??????? IPC??:<br>A01N+A23B+A23C+A23J+A23L+A61K+A61L+A61M+B01D+B01F+B01J+B03C+B05B+B05C+B05D+B07B+B09B+E2<br>2F+B23B+B23C+B23D+B23K+B23Q+B24B+B25J+B32B+B41M+B62C+C01B+C01F+C01G+C02F+C03B+C03C+C04B<br>+C07B+C07C+C07D+C07F+C07H+C07J+C07K+C08B+C08F+C08G+C08J+C08K+C08L+C09C+C09D+C09K+C12N+C<br>12P+C12Q+C21D+C22B+C22C+C23C+C23D+C23F+C23G+C23H+C25B+C25C+C25D+C25F+C30B+B01F+B03D+B04H+B0<br>6F+B06H+B06N+B21H+B01B+B01C+B01J+B01N+B01N033+B01P+B01R+B01T+B02B+B02F+B03C+B03G+B03H+B<br>05D+B06F+B11B+B11C+B12B+B21K+B01B+B01F+B01G+B01J+B01L 021+B01L 023+B01L 025+B01L<br>027+B01L 029+B01L 031+B01L 033+B01L 039+B01L<br>041+B01L049+B01M+B01S+B04B+B05B+B05G+B05H+B05K<br>4) 1+3 | 7,943                      |
| Information             | IPC: G06F+H01L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 49,995                     |
| Environment improvement | Faucet = ZAB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 7,555                      |
| Total                   | -                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 880,043                    |

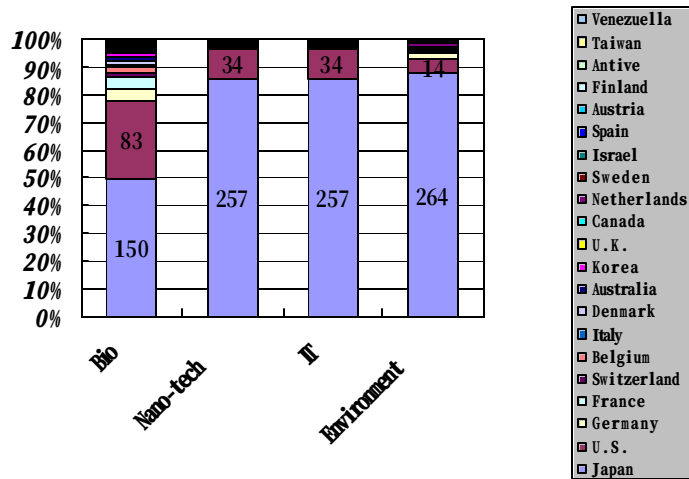
## Citations in each technological fields

|                                | Cited papers |         | Cited patents |         |
|--------------------------------|--------------|---------|---------------|---------|
|                                | total        | Average | Total         | average |
| <b>Biotechnology</b>           | 3,439        | 11.46   | 1,102         | 3.67    |
| <b>Nano-technology</b>         | 597          | 1.99    | 2,125         | 7.08    |
| <b>Information technology</b>  | 95           | 0.32    | 927           | 3.09    |
| <b>Environment improvement</b> | 77           | 0.26    | 1,193         | 3.98    |
| <b>Random sample</b>           | 179          | 0.6     | 1,749         | 5.83    |

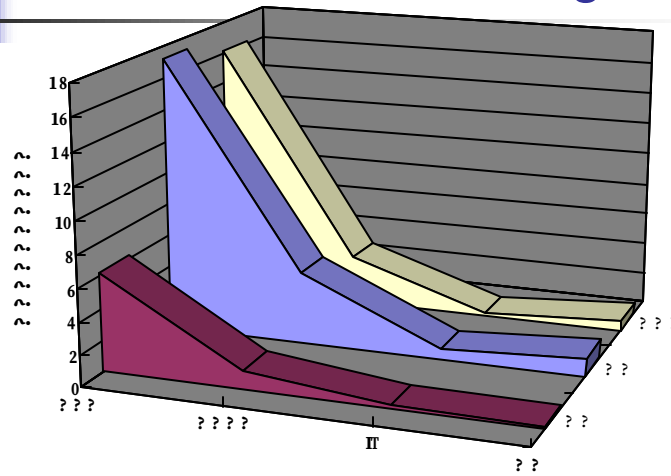
## Number of cited papers in each technological field



## Nationality of assignee



## Science linkage of patents invented in various region



## Preliminary findings

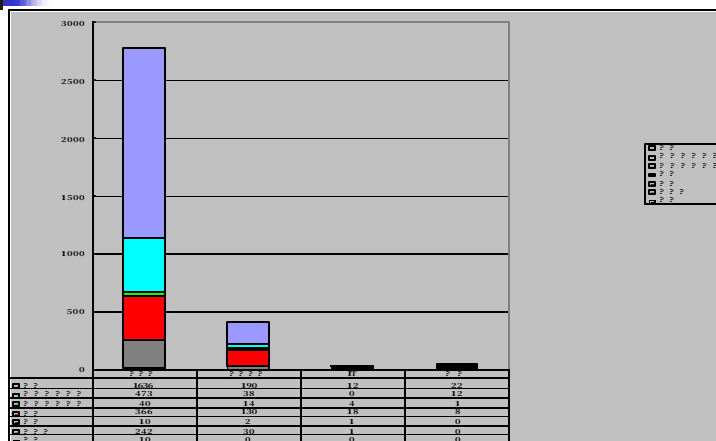
- We constructed an original Japanese patent database and specialized workstation.
- Similar to the U.S. and the EU, inventors cite many academic papers in the Japanese Patent System.
- But they are not in the front page but in the description of the patent gazettes
- Manual extraction of reference cited in patents is labor intensive and time consuming. Therefore, full-automatic extraction of cited reference in the texts of the patent applications is necessary.

## 2. Who supports who's scientific research cited by whose patents

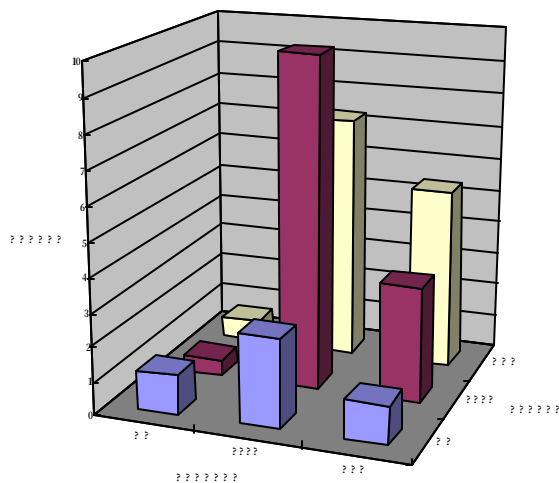
### Accumulation of cited papers



## Characteristics of organization of cited patents



## U.S. papers are cited most even in patents invented in EU and Japan





## Supporting body of biotech papers

|                                                |      |     |
|------------------------------------------------|------|-----|
| Not supported                                  | 1002 | 23% |
| NIH                                            | 547  | 13% |
| NSF                                            | 222  | 5%  |
| NCI (National Cancer Institute)                | 200  | 5%  |
| USPHS (U.S. Public Health Service)             | 168  | 4%  |
| American Cancer Society                        | 157  | 4%  |
| Ministry of science and education, Japan       | 93   | 2%  |
| National Institute of General Medical Sciences | 89   | 2%  |
| Deutsche Forschungsgemeinschaft                | 66   | 2%  |
| U.S. DOE                                       | 57   | 1%  |



## Preliminary findings

- Most of the authors of papers cited in patents belong to academia i.e. universities and public labs.
- U.S. papers are most cited by inventions in Europe and Japan.
- 77% of accumulated papers in biotechnology are publicly funded.



### 3. Full-automatic extraction of cited references

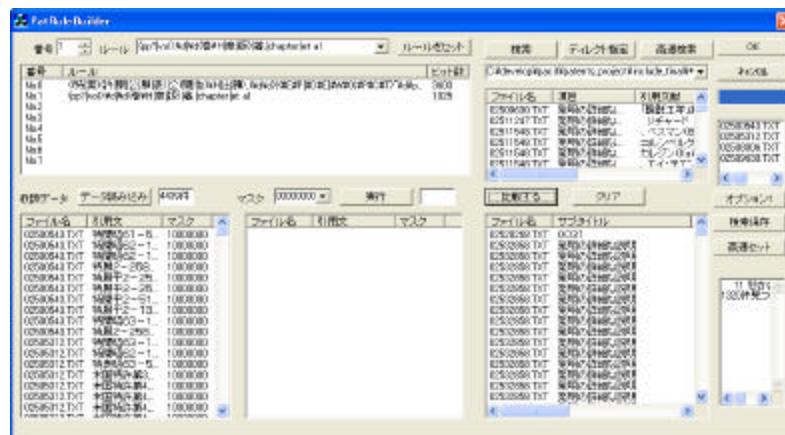


Performance indicator for  
automated reference extraction  
( Recall ratio=R, Precision ratio=P)

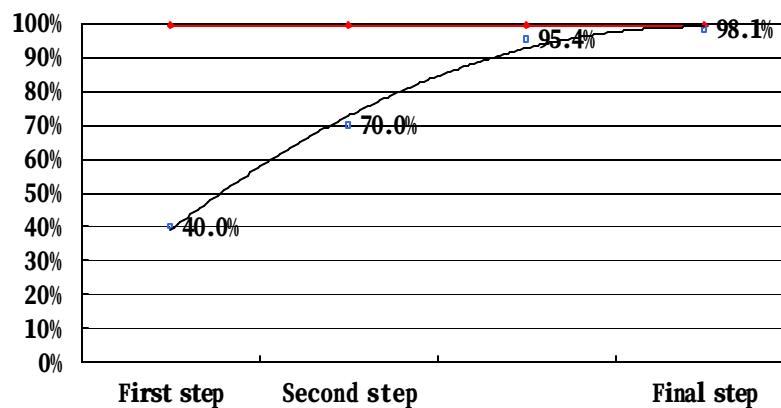
|           | Reference extracted<br>automatically | Reference not extracted<br>automatically |
|-----------|--------------------------------------|------------------------------------------|
| Matched   | w                                    | x (leakage)                              |
| Unmatched | y (noise)                            | z                                        |

$$R = \frac{w}{w + x} \quad P = \frac{w}{w + y}$$

## Improvement of search formula



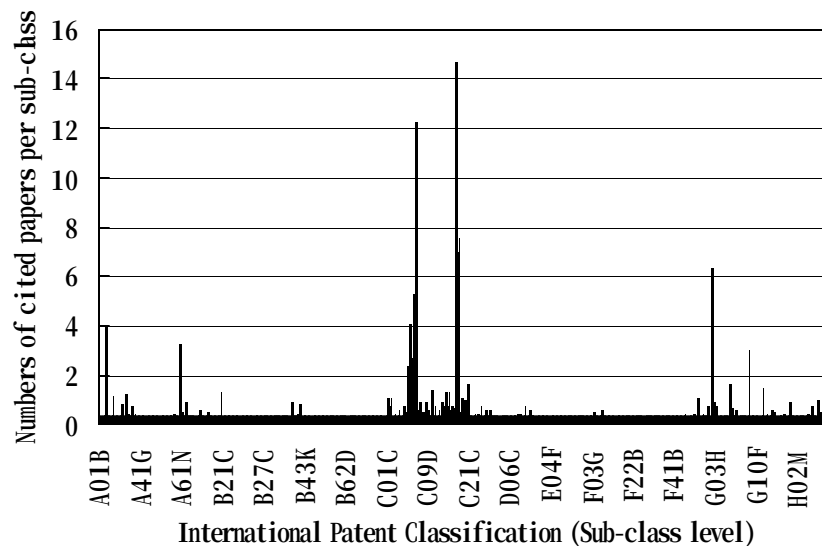
## Recall ratio increased by program improvement



## Calculation result of automated reference extraction

|                  | patent reference |       |         | non-patent reference |       |         |
|------------------|------------------|-------|---------|----------------------|-------|---------|
| Categories       | patenet cited    | noise | leakage | non-patent cited     | noise | leakage |
| Env-tech unexam  | 531              | 1     | 0       | 55                   | 2     | 11      |
| Env-tech exam    | 1296             | 0     | 5       | 73                   | 0     | 3       |
| Mixed unexam     | 1355             | 7     | 6       | 435                  | 6     | 18      |
| Mixed exam       | 2342             | 14    | 2       | 672                  | 11    | 32      |
| IT-tech unexam   | 234              | 2     | 4       | 46                   | 8     | 2       |
| IT-tech exam     | 977              | 5     | 6       | 115                  | 2     | 18      |
| Bio-tech unexam  | 875              | 14    | 6       | 3420                 | 112   | 31      |
| Bio-tech exam    | 1322             | 25    | 13      | 4267                 | 22    | 32      |
| Nano-tech unexam | 476              | 4     | 2       | 83                   | 11    | 11      |
| Nano-tech exam   | 1867             | 1     | 1       | 213                  | 3     | 14      |
| SUM              | 11275            | 73    | 45      | 9379                 | 177   | 172     |
|                  | Recall ratio     |       | 99.6%   | Recall ratio         |       | 98.2%   |
|                  | Precision ratio  |       | 99.4%   | Precision ratio      |       | 98.1%   |

## Significant difference in science linkage among different technological classifications



## Ranking of Technological Fields (top 10)

-Which technological field represents strong science linkage in Japan?-

| Sub-classes                                                                                                                                                                                                                          | Number of Patents | Science Linkage |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|-----------------|
| C12N MICRO-ORGANISMS OR ENZYMES; COMPOSITIONS THEREOF; PROPAGATING, PRESERVING, OR MAINTAINING MICRO-ORGANISMS; MUTATION OR GENETIC ENGINEERING; CULTURE MEDIA                                                                       | 44425             | 14.6            |
| C07K PEPTIDES                                                                                                                                                                                                                        | 18390             | 12.3            |
| C12Q MEASURING OR TESTING PROCESSES INVOLVING ENZYMES OR MICRO-ORGANISMS; COMPOSITIONS OR TEST PAPERS THEREFOR; PROCESSES OF PREPARING SUCH COMPOSITIONS; CONDITION-RESPONSIVE CONTROL IN MICROBIOLOGICAL OR ENZYMOLOGICAL PROCESSES | 5442              | 7.6             |
| C12P FERMENTATION OR ENZYME-USING PROCESSES TO SYNTHESISE A DESIRED CHEMICAL COMPOUND OR COMPOSITION OR TO SEPARATE OPTICAL ISOMERS FROM A RACEMIC MIXTURE                                                                           | 9617              | 7.0             |
| G03C PHOTSENSITIVE MATERIALS FOR PHOTOGRAPHIC PURPOSES; PHOTOGRAPHIC PROCESSES, e.g. CINE, X-RAY, COLOUR, STEREO-PHOTOGRAPHIC PROCESSES; AUXILIARY PROCESSES IN PHOTOGRAPHY                                                          | 24018             | 6.3             |
| C07J STEROIDS                                                                                                                                                                                                                        | 1373              | 5.3             |
| C07H SUGARS; DERIVATIVES THEREOF; NUCLEOSIDES; NUCLEOTIDES; NUCLEIC ACIDS                                                                                                                                                            | 2837              | 5.0             |
| C07D HETEROCYCLIC COMPOUNDS                                                                                                                                                                                                          | 24241             | 4.1             |
| A01H NEW PLANTS OR PROCESSES FOR OBTAINING THEM; PLANT REPRODUCTION BY TISSUE CULTURE TECHNIQUES                                                                                                                                     | 596               | 4.0             |
| A61K PREPARATIONS FOR MEDICAL, DENTAL, OR TOILET PURPOSES                                                                                                                                                                            | 23852             | 3.3             |

## Ranking of Technological Fields (11~ 20)

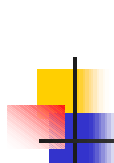
-Which technological field represents strong science linkage in Japan?-

| Sub-classes                                                                                                                                                                                                                             | Number of Patents | Science Linkage |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|-----------------|
| G09C CIPHERING OR DECIPHERING APPARATUS FOR CRYPTOGRAPHIC OR OTHER PURPOSES INVOLVING THE NEED FOR SECRECY                                                                                                                              | 233               | 3.0             |
| C07C ORGANIC COMPOUNDS OF UNKNOWN CONSTITUTION                                                                                                                                                                                          | 198               | 2.7             |
| C07F ACYCLIC, CARBOCYCLIC, OR HETEROCYCLIC COMPOUNDS CONTAINING ELEMENTS OTHER THAN CARBON, HYDROGEN, HALOGEN, OXYGEN, NITROGEN, SULFUR, SELENIUM, OR TELLURIUM (metal-containing porphyrins)                                           | 3651              | 2.6             |
| C08B POLYSACCHARIDES; DERIVATIVES THEREOF (polysaccharides containing less than six saccharide radicals attached to each other by glycosidic linkages; fermentation or enzyme-using processes; sugar industry; production of cellulose) | 1155              | 2.6             |
| C07B GENERAL METHODS OF ORGANIC CHEMISTRY; APPARATUS THEREFOR                                                                                                                                                                           | 468               | 2.3             |
| C07C ACYCLIC OR CARBOCYCLIC COMPOUNDS                                                                                                                                                                                                   | 15291             | 2.0             |
| C14C CHEMICAL TREATMENT OF HIDES, SKINS OR LEATHER, e.g. TANNING, IMPREGNATING, FINISHING; APPARATUS THEREFOR, COMPOSITIONS FOR TANNING                                                                                                 | 51                | 1.6             |
| G06E OPTICAL COMPUTING DEVICES                                                                                                                                                                                                          | 56                | 1.6             |
| G10L SPEECH ANALYSIS OR SYNTHESIS; SPEECH RECOGNITION                                                                                                                                                                                   | 1761              | 1.5             |
| C09H PREPARATION OF GLUE OR GELATINE                                                                                                                                                                                                    | 18                | 1.4             |



## Major findings & conclusion

- Similar to the U.S. and the EU, inventors cite many academic papers in the texts of the patent applications in the Japanese Patent System.
- Full-automatic extraction of cited reference in the texts of the patent applications is accomplished.
- The science linkage index among different patent classifications differs significantly.
- The pattern of difference of science linkage index was similar in the EU and Japan.
- It suggests that the process of creating new technology differs from technology to technology.



## Useful conceptual frameworks & policy recommendations (new)

- Information vs. entropy
  - To require inventors and applicants to provide precise information before putting them into the database is much costless than to clean the information up after having put wrong information and clean them all
- Patent vs. (Privacy & Trade Secret)  
=Public Interest vs. Private Interest
  - If a policy that require applicants to provide additional information accelerate the speed of innovation, that policy should be put into practice.
  - Like JPO applicant code, introducing universal system of applicant code and inventor code is a good policy example.