

研究レビュー 4-1

論文・特許データの統合的データ体系の構築

～データの名寄せの挑戦～

科学技術基盤調査研究室 富澤 宏之

講演

特許情報を使った科学技術イノベーション政策研究の事例報告

一橋大学 長岡 貞男

研究レビュー 4-2

NISTEPにおける特許情報関連調査研究のこれまでとこれから

第3調査研究グループ 細野 光章

Science and Technology Policy Review Vol.4

March 2013

National Institute of Science and Technology Policy(NISTEP)
Ministry of Education, Culture, Sports, Science and Technology (MEXT)
Japan

本報告書の引用を行う際には、出典を明記願います

目 次

科学技術政策研究レビューの趣旨	i
-----------------------	---

〔研究レビュー 4-1〕

論文・特許データの統合的データ体系の構築：データの名寄せの挑戦

はじめに	1
1 科学論文データによる名寄せ	3
2 特許データにおける名寄せ	21
3 データ整備がもたらすもの：増大するデータの活用の試み	24

〔講演〕

特許情報を使った科学技術イノベーション政策研究の事例報告

はじめに	27
1 特許情報の基本的内容	28
2 発明の企業秘密による保護と特許出願	31
3 発明者と発明チーム	32
4 国際共同発明	34
5 発明者の生産性	35
6 知識ストックと知識スピルオーバー	36
7 おわりに	38

〔研究レビュー 4-2〕

NISTEP における特許情報関連調査研究のこれまでとこれから

はじめに	41
1 科学技術・イノベーション政策研究の対象としての特許情報（制度）	42
2 NISTEP におけるこれまでの特許情報（制度）の調査研究	43
3 NISTEP における特許情報（制度）の調査研究の現状とこれから	48

科学技術政策研究レビューの 趣旨

科学技術政策研究レビューの刊行に当たって

第4期科学技術基本計画においては、科学技術イノベーション政策をいかに実効あるものにしていくかが重要な課題となっております。このような政策形成に当たってのさまざまなエビデンスを提供することは当研究所の使命であり、多様な研究活動を進めております。

最新のデータ等を関係する行政部局等にできるだけ早く提供するという観点から、ひとつの調査研究が終了すると、その成果を単発のレポートとして取りまとめています。その結果として、科学技術政策に関する大きなテーマについて、調査案件毎に細分化されたレポートが独立に存在しており、科学技術政策研究所の調査研究活動全体として何が見えているのか、何が大きな課題なのかというような俯瞰が充分説明できていないのではないかという問題意識を持つようになりました。

そこで、2011年度より、科学技術政策研究レビューを発行し、ある程度大きなテーマについて当研究所の研究成果を中心とする俯瞰的レビューを行うこととしました。執筆者は、担当テーマについての政策の流れ、内外の政策研究の動向、他のテーマとの関連性等についての考察にも取り組みます。このような活動は、次に取り組むべき研究課題を浮き彫りにするための「マッピング」としての機能も持つものであり、様々な関係者の皆様からご意見をいただくことも重要と考えております。

科学技術政策研究レビューは当面年2回程度発行していく方針で、本誌はその第4号にあたります。今回は、(1)論文・特許データの整合的データ体系の構築：データの名寄せの挑戦、(2)特許情報を使った科学技術イノベーション政策研究の事例報告、(3)NISTEPにおける特許情報関連調査研究のこれまでとこれから、の3つのテーマを取り上げています。

最後になりましたが、私ども科学技術政策研究所の調査研究活動につきまして、今後ともご指導、ご鞭撻をいただくことをお願い申し上げます。

2013年3月
科学技術政策研究所
所長 桑原 輝隆

〔研究レビュー 4-1〕

論文・特許データの整合的データ
体系の構築：データの名寄せの挑戦

科学技術基盤調査研究室 富澤 宏之

研究レビュー 4-1

特許データの統合的データ体系の構築:データの名寄せの挑戦

科学技術基盤調査研究室 富澤 宏之

はじめに

本報告は、標題が示すように、データの分析結果ではなく、データの整備という技術的・方法論的な内容についての報告であり、科学技術政策に直接的な示唆を与えるものではない。しかし、質の高いデータの整備は、有用な分析が可能かどうかを左右する決定的要素であるため、その一端を紹介するとともに、今後の方向性についての考え方を示し、データ整備や活用についての議論を活発化させることが狙いである。なお、本報告は、文部科学省が 2011 年度より実施している「科学技術イノベーションにおける政策のための科学」推進事業の一環として科学技術政策研究所が取り組んでいる「データ・情報基盤の構築」及び関連する取り組みの一部について述べたものである。

本報告は、データの名寄せがテーマであるが、その内容を述べるに先立ち、データの名寄せは特殊な技術的問題ではなく、現代社会においてしばしば顕在化する問題であることを指摘したい【資料 1】。例えば、2007 年に社会保険庁の年金記録の不備が大きな社会的問題となった際には、年金記録データについての「名寄せが必要」といった表現が広く用いられた。これは、コンピューターに入力された年金記録データのなかに、同一人物の重複や入力ミスなどが大量に含まれているという問題であり、本報告で取り上げる「名寄せ」と同列に語るべきではないが、適切に作成されていないデータが大きな問題を引き起こすことがあることをよく示している。

また、社会科学において、データクリーニングがなされていない調査データは、ほとんど価値が無いに等しいという認識が研究者や専門家の間で広く共有されており、データの名寄せはデータクリーニングのなかでも中心的課題となっている。情報科学においても、データマイニング作業の大部分はデータクリーニングである、という意見がある。

【資料1】 データの名寄せ/データクリーニングの意義

- データの名寄せ/データクリーニングの問題の広さ
 - 年金記録問題...
 - 社会調査:「クリーニングのなされていない調査データは無価値」
 - 情報科学:「データマイニング作業の大部分はデータクリーニング」

- より積極的な意義
 - “情報爆発”の時代の重要な技術?
 - 多様な情報を活用した様々な定量的研究が可能に
 - 雑多なデータやノイズの増大
 - 分析目的で作成されていないデータの定量分析への転用
 - 例:科学計量学/特許統計(論文や特許の検索性DBの活用)
 - 例:データマイニング
 - 単なるデータの塊(非構造化データ)から構造化データへ
 - 異種データの連結

データの名寄せやクリーニングは、必要に迫られて行う作業というだけでなく、より積極的な意義もある。“情報爆発”という現象が語られる今日において、社会や経済の諸活動を通じて蓄積された膨大なデータが、その名寄せやクリーニングにより、新たな価値を生み出す可能性があるためである。研究の世界においても、本来、分析目的で作成されていないデータを定量的研究に用いようとする試みが、今後、一層、盛んになると考えられる。そもそも、科学計量学、つまり、本報告でとりあげる科学論文や特許の定量的な分析自体も、本来、論文や特許の検索のために作成されたデータベースを定量的研究という別の目的に転用した例と言えるのではないだろうか。このようなデータの活用が盛んになると、単なる個別情報の集まりに過ぎないデータセットをいかに構造化するか、といったことが重要になる。また、異なるデータをどのように連結するか、という点も重要になってくる。

1. 科学論文データにおける名寄せ

1.1 科学論文データの機関レベルでの名寄せ

データの名寄せについての具体的な議論として、本報告では主に、科学論文データにおける名寄せについて述べる。

まず、名寄せを行う前の科学論文データの例を【資料 2】に示した。これは、科学論文の代表的なデータベースである Web of Science(WoS)の1レコードを示している。これを見ると、「著者所属」の部分には、MIT やブラジルの大学、さらに日本の東北大学の名称とアドレスが記載されている。これらは、上方に記載された著者名と対応づけられている。また、「助成金」の部分には、米国やブラジルの研究ファンドとともに、MEXT、すなわち日本の文部科学省についての記載があり、「20241023」という番号が示されている。これは科学研究費補助金（現在は科学研究費助成制度）の研究課題番号である。このような情報は、元の論文においては謝辞（acknowledgement）の中に書かれたものであり、それがこのようにデータベースに収録されるようになったのは、ごく最近のことである。このような情報により、それぞれの論文がどのような研究費によって実施された研究の成果であるかを知ることができる。ただし、論文の謝辞の書き方は統一されていない場合が多く、現時点において、データの質は高くはない。

【資料 2】 科学論文データベースにおけるデータの例

Perspectives on Carbon Nanotubes and Graphene Raman Spectroscopy

著者名: Dresselhaus, MS (Dresselhaus, Mildred S)^{1,2}, Jorio, A (Jorio, Ado)³, Hofmann, M (Hofmann, Mario)⁴, Dresselhaus, G (Dresselhaus, Gene)⁴, Saito, R (Saito, RIchiro)⁵

出版物名: NANO LETTERS 巻: 10 号: 3 ページ: 751-758 DOI: 10.1021/nl904286f 発行: MAR 2010

被引用数: 187 (Web of Science から)

引用文献: 38 [Related Records を検索] [引用マップ](#)

抄録: Raman spectroscopy is here shown to provide a powerful tool to differentiate between two different sp² carbon nanostructures (carbon nanotubes and graphene) which have many properties in common and others that differ. Emphasis is given to the richness of both carbon nanostructures as prototype examples of nanostructured materials. A glimpse toward future developments in this held is presented.

アクション番号: WOS:000275278200001

ドキュメントタイプ: Article

言語: English

著者によるキーワード: Graphene; carbon nanotubes; graphene ribbons; Raman spectroscopy

KeyWords Plus: OPTICAL-TRANSITION ENERGIES; SCATTERING; GRAPHITE

別刷り請求先: Dresselhaus, MS (別刷り著者), MIT, Dept Elect Engrn & Comp Sci, Cambridge, MA 02139 USA

著者所属:

1. MIT, Dept Elect Engrn & Comp Sci, Cambridge, MA 02139 USA
2. MIT, Dept Phys, Cambridge, MA 02139 USA
3. Univ Fed Minas Gerais, Dept Phys, BR-30123970 Belo Horizonte, MG, Brazil
4. MIT, Francis Bitter Natl Magnet Lab, Cambridge, MA 02139 USA
5. Tohoku Univ, Dept Phys, Sendai, Miyagi 9808578, Japan

Email アドレス: millie@mgm.mit.edu

ResearcherID 番号: [?]

[2 人の研究者の ResearcherID の [My Publication] リストに、このレコードが含まれています。クリックして表示。]

助成金:

助成金提供機関	助成金登録番号
MCT-CNPq (Brazil)	
AFOSR (USA)	
MEXT	20241023
	NSF-DMR-07-04197

[\[助成金情報を表示\]](#)

発行者: AMER CHEMICAL SOC, 1155 16TH ST, NW, WASHINGTON, DC 20036 USA

Web of Science の分野: Chemistry, Multidisciplinary; Chemistry, Physical; Nanoscience & Nanotechnology; Materials Science, Multidisciplinary; Physics, Applied; Physics, Condensed Matter

主観分野: Chemistry, Science & Technology - Other Topics; Materials Science; Physics

IDS番号: 565GR

ISSN: 1530-6984

研究レビュー4-1

特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

次に、科学論文データベースの主なデータ項目（フィールド）を【資料3】に示した。一見すると、データ項目数は多くないため、豊富な情報には見えないかもしれない。しかし、各論文に関する基本的な書誌情報がよく整理されており、それが数千万件について集積されているため、統計データとしての価値が高く、様々な角度からの集計・分析が可能である。特に、定量分析が可能であるための基本的要件であるデータの一意性（1レコードが1論文に対応）が確保されていることが重要である。そして、それを超えて更に豊富な情報を引き出すための有力な手段の一つがデータの名寄せなのである。

【資料3】 科学計量学で用いられる科学論文データベースの主なデータ項目

- Thomson Reuters, “Web of Science”の主要なデータ項目
 - 出版物名(ジャーナル名や書籍名)
 - 出版年
 - 論文タイトル
 - DOI (デジタルオブジェクト識別子=論文・文献のID)
 - 抄録
 - 著者名
 - 著者所属
 - 言語
 - ドキュメントタイプ(アーティクル、レビュー、ノート、書籍、などの区別)
 - 会議情報
 - 助成金提供機関
 - 助成金登録番号
 - 参照文献
 - 当該論文を引用している文献(→被引用回数)

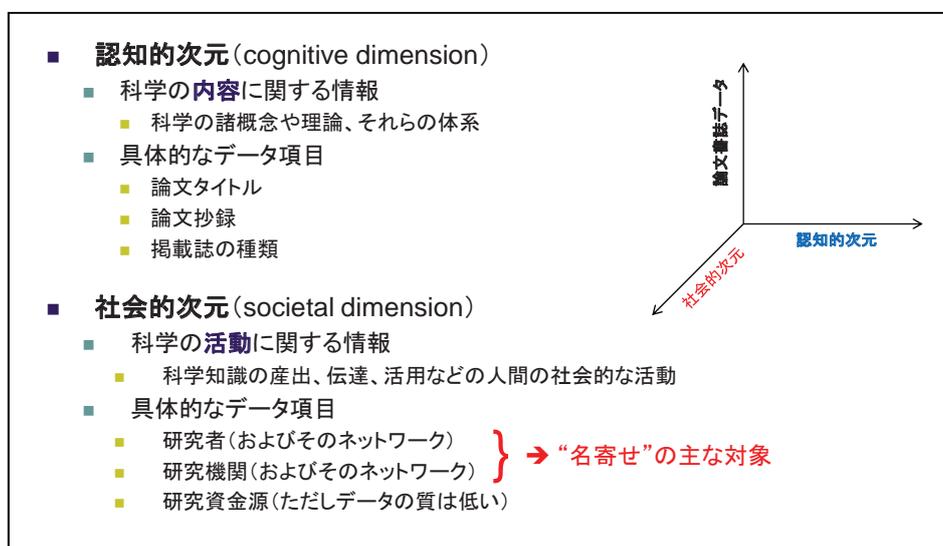
ところで、【資料3】に示したような科学論文データベースのデータ項目は、大きく2つに分けることができる【資料4】。第一は、「科学の内容」に関するデータである。すなわち、論文のタイトルや論文のアブストラクトなどであり、その論文がどのような科学研究を扱っているのかという情報を含んでいる。そのようなデータを分析することにより、例えば、最近、どのような科学研究が盛んであるのか、ホットな研究領域はどれか、といったことを明らかにすることができる。これは、人間の思考や知的活動の内容に関わる情報であるため、認知的次元(cognitive dimension)と呼ぶことができよう。

特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

もう一つは、科学の社会的な活動としての側面についての情報であり、社会的次元 (societal dimension) と呼ぶことができる。具体的なデータ項目としては研究者 (共著者も含む) の名前、研究者が所属する研究機関の名称や所在地、研究を実施するための研究資金源など、科学活動に関する情報である。本報告が主題としている名寄せとは、主に、この社会的次元に関するデータを対象としたものである。先に見たようなデータ項目のある部分を整備・構造化していくことによって、この社会的次元に関する部分が浮かび上がってくるのである。

【資料 4】 科学計量学の対象となるデータの2つの次元



次に、科学論文データベースについての名寄せがどのようなものであるのか、具体的に述べたい。例として、ある特定の機関の論文発表数を集計したい場合を考える。データベースには、論文に記載された機関名 (著者の所属機関名) がそのまま収録されているのが普通であるが、類似の名称の機関の存在、名称表記の揺らぎ、あるいは誤記などがあるため、データベースに単純な検索をかけても正確な件数が得られるとは限らない。また、機関の表記のレベルについてのばらつき、つまり、大学全体を表記する場合や、大学の付属の研究所名のみを表記する場合、といった点についてのばらつきもある。さらには、機関の統廃合や名称変更もあるため、単純な検索では、求めるデータが得られない場合がある。

特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

【資料 5】の右側には、機関名称の揺らぎの例を示した。3つの大学を取り上げているが、これらは実際に WoS データベースの中に収録された大学名称を抽出したものである。このように名称の揺らぎがあるため、これらを欠落なく全て拾い上げるのが重要な課題となる。

さらに問題であるのは、ただ単に正確でないということだけではなく、そのデータの精度がどのくらいかよくわからないことである。そのようなデータの精度の評価がなされていないと、そのデータを用いた分析の精度にも疑問が生じることになる。

【資料 5】 科学論文データベースによる機関別データの作成の困難

特定の機関の論文数を 得たい場合・・・	機関名称の揺らぎの例 (WoS収録データ)	
<ul style="list-style-type: none"> ■ 論文データベースには、原則的に、論文に記載された機関名がそのまま収録 ■ 科学論文データベースの単純な検索では、正確な件数は得られない <ul style="list-style-type: none"> ■ 類似名称の機関の存在 ■ 名称表記の揺らぎ・誤記の存在 ■ 機関の階層構造のばらつき ■ 機関の統廃合 ■ 得られた数値データの精度を評価できない 	DENKI TSUSHIN UNIV	電気通信大学
	UNIV ELECTRO COMMUN	電気通信大学
	UNIV ELECTROCOMMUN	電気通信大学
	UNIV ELETROCOMMUN	電気通信大学
	NAIST	奈良先端科学技術大学院大学
	NARA ADV INST SCI & TECHNOL	奈良先端科学技術大学院大学
	NARA INST SCI & TECHNOL	奈良先端科学技術大学院大学
	NARA INST SCI & TECHNOL 89165	奈良先端科学技術大学院大学
	NARA INST SCI & TECHNOL NAIST	奈良先端科学技術大学院大学
	NARA INST SCI & TECHOL	奈良先端科学技術大学院大学
	NARA INST SCI TECHNOL	奈良先端科学技術大学院大学
	OSAKA UNIV	大阪大学
	OSAKA UNIV FOREIGN STUDIES	大阪大学
	OSAKA UNIV HOSP	大阪大学
UNIV OSAKA	大阪大学	

また、名称の揺らぎだけでなく、紛らわしい類似の名称を持つ異機関もあり、【資料 6】では2つの例を挙げている。【資料 6】の左側の例は、大阪大学とそれに類似する名称を持った機関であり、後者は、「Osaka University」という単語を含んでいるものの、これは財団法人で大阪大学とは異なる組織で、所在地も香川県である。このように、名称だけでは判別が難しいものが含まれている。

特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

【資料 6】の右側は、更に深刻な例である。これは、国立の静岡大学と静岡県立大学の例であり、両者は日本語でも英語でも名称が似ている上に、所在地情報についても「静岡県静岡市駿河区」という部分まで共通している。それでも、この図にある通りの正式名称が科学論文データベースに記載されているならば判別は可能であるが、静岡県立大学の英語名が「Shizuoka University」と誤った表記がされている、あるいはその逆といった場合があり、名称だけでは正しく区別できない場合がある。そのようなデータについては、詳しく所在地情報を参照する、あるいは両大学の学部構成まで考慮する、といった手間をかけてデータを整理することが必要となる。

【資料 6】 紛らわしい名称を持つ機関の例



このように、少数の機関の場合でも正確なデータを作成するためには多大な手間を要するため、例えば、10~20程度の主要大学に限って分析する方法が考えられるが、その場合でも、データを作成するための手間は多大である。例えば、一つの大学でも、どのような附属研究機関や学部があるか、等について調べる必要があるため、10~20大学を調べる場合でも、実際には数百の関連する部局を調べる必要が生じる可能性がある。そのため、ある程度の数の大学や機関についてのデータを整備しようとするのであれば、逆説的であるが、日本の全ての機関について機関別データを整備したほうが、むしろ効率的であると考えられる。

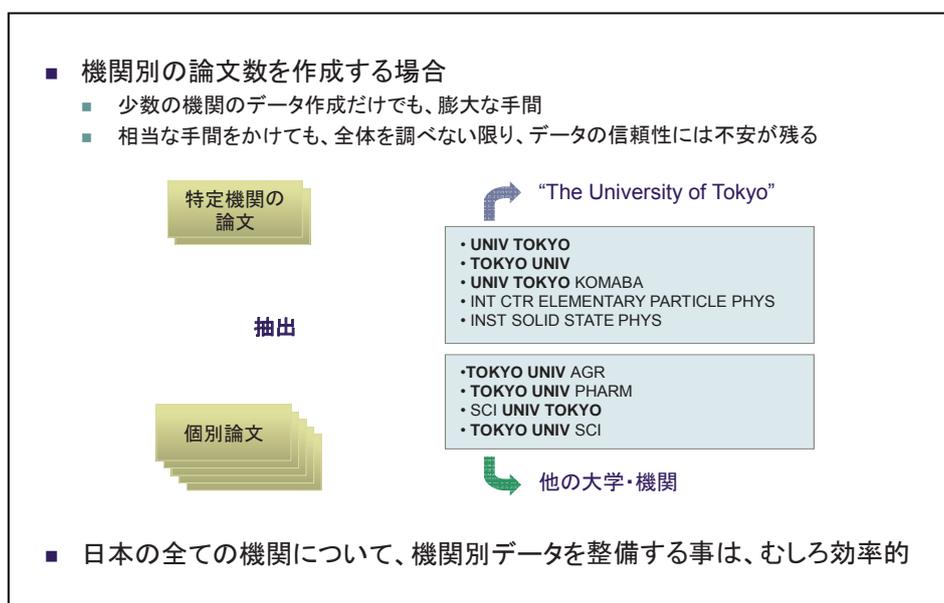
研究レビュー4-1

特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

【資料7】に示すように、例えば、東京大学（The University of Tokyo）についてのデータを得たい場合、「Tokyo Univ Agri」（東京農工大学）や「Tokyo Univ Pharm」（東京薬科大学）、「Sci Univ Tokyo」（東京理科大学）などの類似する名称の大学と区別する必要があるが、もし、別途、これらの大学についても名寄せしているのであれば、この区別は、比較的容易になる。つまり、単に「対象 a が X でないこと」を確認しようとするのはそれほど容易でないが、「対象 a が Y であること」を確認できれば、「対象 a が X でないこと」は確実に判定できるということである。このような方法は、全体として効率的である上に、データの信頼性を確保する意味でも利点がある。

【資料7】 科学論文データベースの機関名の統合的な名寄せの意義(1)

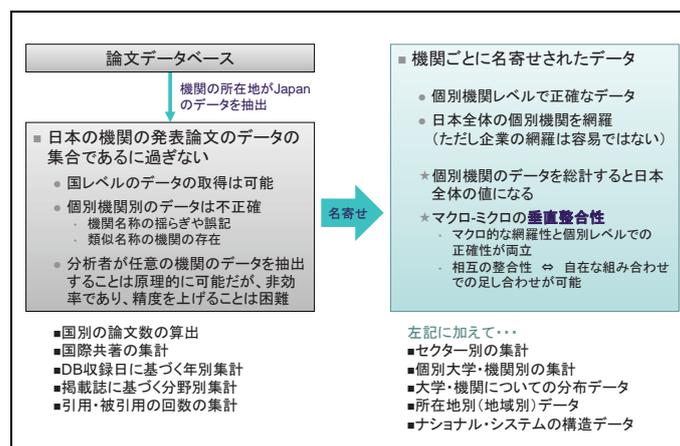


特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

【資料 8】では、データの名寄せの意義を示すために、名寄せ前のデータと名寄せ後のデータがどのように異なるのかを比較している。図の左側にあるように、科学論文データベースを用いて日本について分析したい場合、機関の所在地に「Japan」という語が含まれるデータを抽出することが出発点となる。これにより、日本全体の論文の集合ができるため、日本全体のマクロ的な論文数や、日本と外国との国際共著論文数や被引用回数の集計、あるいはそれらの時系列データの取得も可能になる。また、それらの分野別の集計や、被引用度の高い論文数の集計といった集計も可能である。このような集計に基づくデータは、科学技術指標として世界各国で使われている。しかし、このようなマクロ集計値を超えて、より詳細な定量データを得ることは、困難である。

【資料 8】 科学論文データベースの機関名の統合的な名寄せの意義(2)



一方、図の右側は、名寄せを行ったデータの利点を示している。名寄せを行うことにより、個別機関レベルで正確なデータが得られ、しかも、日本全体の個別機関が網羅されたデータも原理的には取得可能である。それに加えて、個別機関のデータを総計すると日本全体の値になるという、マクロレベルのデータとミクロレベルのデータとの整合性が確保されることも可能になる。このことは、個別機関の自由な組み合わせの足し合わせが可能になることを意味している。例えば、医学部を有する大学全体の論文数や、それが日本の大学全体の中でどの程度の割合を占めるのか、あるいは、九州に所在する大学の論文数の合計、などのような様々な集計が可能になる。さらには、機関別の論文数の分布のデータといった分析も可能となり、ナショナルシステムの構造データとしての意味を有するようになる。

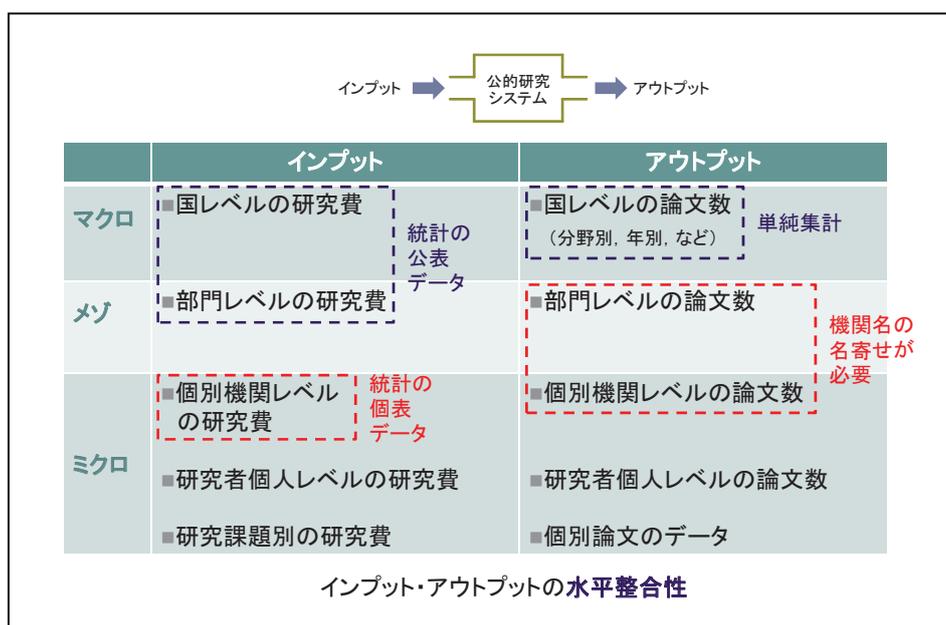
研究レビュー4-1

特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

今述べたような、マクロデータとマイクロデータとの間の整合性は、垂直整合性と呼ぶことができるが、さらに、水平整合性、すなわちインプットとアウトプットのデータのような異種のデータ間の整合性を得るためにも、データの名寄せが決定的に重要となる【資料9】。なぜなら、インプットとアウトプットの両方で同じレベルのデータが揃うことにより、はじめて両データ間の接続が可能となるためである。

【資料9】 科学論文データベースの機関名統合的な名寄せの意義(3)



なお、ここで言うインプットとは、公的研究機関に投入される研究資源（特に研究費）が代表的なものであり、一方、アウトプットとは研究成果であり、その代表的なものは科学論文である。インプット側で入手可能なデータは、最近まで、統計の公表データにほとんど限定されていたが、これらは、マクロレベルやメゾレベル（部門レベルのような中間レベル）のデータに限られていた。しかし、最近、統計の個票データを研究目的で使うことができるようになり、個別機関レベルの研究費といったようなデータが用いられるようになってきている。

一方、アウトプット側については、単純集計により国レベルの論文データを得ることは比較的容易であり、よく用いられてきたが、さらに機関名の名寄せを行うことにより、セクターや個別機関のレベルの論文データが得ることができるようになる。そうするとインプットとアウトプットのデータが各レベルで揃うことになるため、両者の連結が統合的に可能になる。

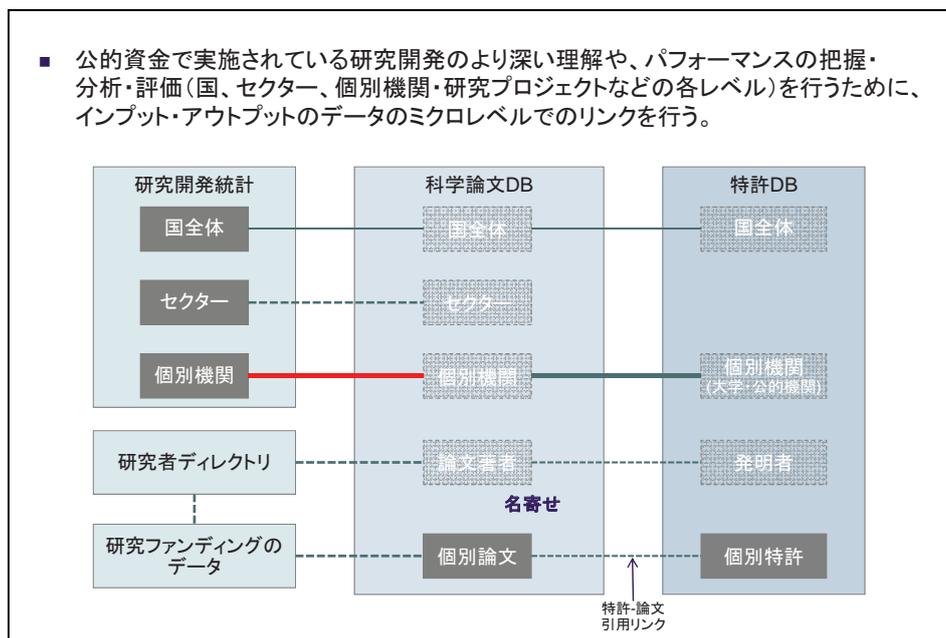
特許データの整合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

1.2 科学技術政策研究所の取り組み

以上のような考えに基づき、科学技術政策研究所では、文部科学省の科学技術イノベーション政策における「政策のための科学」推進事業の一環として、「データ・情報基盤の構築」を実施している。そのなかのプロジェクトの一つとして、「大学・公的研究開発機関に関するデータ整備」に取り組んでおり、【資料10】のような概念モデルに基づいてデータ整備を行っている。すなわち、研究開発統計と科学論文のデータベース、さらに部分的に特許のデータベースを用い、それらを様々なレベルで整備して、相互に連結しようとするものである。この図には、このデータの整備において、データの名寄せがどのような位置づけにあるのかを示している。すなわち、データの名寄せがなされていない場合、データ接続は、個別論文レベルか、あるいは国全体やセクターのレベルでのみ可能であるが、名寄せにより論文著者（個人）や個別機関のレベルのデータが作成されることにより、それぞれのレベルで他のデータとの接続が可能になるのである。

【資料10】 大学・公的研究開発機関に関するデータ整備の概念モデル



研究レビュー4-1

特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

この「公的研究開発機関に関するデータ整備」のうち、初年度にあたる 2011 年度のデータ整備の概要を【資料 11】に示した。まず、日本の大学と公的機関を主な対象とした機関名辞書の作成を行った。これは、様々なデータを個別機関レベルで連結するための基礎であり、全体の中核的存在となるものである。第二に、Scopus (科学論文データベース) のデータについて、機関名の名寄せを実施し、機関名辞書と連結できるようにデータを整備した。第三に、機関名辞書を各種統計データとリンクさせるための情報の整理を行った。この情報は、各種の個別機関レベルのデータを Scopus の機関名寄せ後のデータと連結させるためのツールとして用いることが可能である。第四に、論文著者別のデータを整備するための基礎として、個人識別のアルゴリズムの予備的検討を行った。科学論文データベースに収録されている論文著者情報は、同姓同名や所属組織の変更などがあるなど、個人を完全に識別するために十分な情報ではないため、ここで検討したアルゴリズムは、一定の精度の範囲でのみ有効であるに過ぎない。しかし、論文著者別のデータは、科学研究の基本的な実施主体である研究者個人に着目した研究を行うためには、極めて重要である。

【資料 11】 大学・公的研究開発機関に関するデータ整備：2011 年度のデータ整備の概要

<ul style="list-style-type: none">■ 機関名辞書の作成 ①<ul style="list-style-type: none">■ 基本的に、公開情報に基づいて作成■ 機関名辞書に基づくScopusデータの名寄せ ②<ul style="list-style-type: none">■ 日本の大学等と公的機関 (2000~2010年)<ul style="list-style-type: none">■ 大規模大学(12大学)については、さらに、部局(学部や研究科等)レベルで名寄せ■ 個別機関が同定されないものについても「大学、政府研究機関、会社、非営利団体、その他」などのセクター分類■ 「病院」にはフラグを付与■ 達成目標:「不明が2%以下」	<ul style="list-style-type: none">■ 機関名辞書と各種統計等とのリンク<ul style="list-style-type: none">■ 「科学技術研究調査」の調査対象名簿 (2002~2010年; 大学と公的機関) ③■ 特許公報DB (2002~2010年; 大学と公的機関) ④■ PATSTATの非特許文献とScopus (1996~2010年) ⑤■ KAKENとScopus(1996~2010年) ⑥■ 著者データの整備<ul style="list-style-type: none">■ 個人識別のアルゴリズムの予備的検討<ul style="list-style-type: none">■ 論文DB中の氏名と共著者
--	---

特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

次に、全体の中核的存在である機関名辞書の整備状況を述べる。【資料 12】の左側の表は、機関名辞書における収録機関数を機関種別ごとに示している。総数約 1 万機関の機関が登録されており、各機関については、機関名称（日本語と英語）に加えて所在地や機関種別などの情報を含んでいる。また、大学の統廃合などの情報も収録している。また、大規模な 12 の国立大学のみについては、部局のレベルでの名寄せも実施した。右側の表はその状況を示している。例えば北海道大学の場合、大学の学部や研究所など、69 部局が収録されている。科学論文における機関名の表記に基づいているため、実際の部局数以上に詳細に分類されている可能性があるが、論文データベースに基づいて整理した場合の結果を示している。

【資料 12】 NISTEP 大学・公的機関名辞書の整備状況

機関の種別ごとの登録機関数		部局レベルでの名寄せの結果 (2011年度は12の国立大学を対象)	
機関の種別	登録機関数	機関名	部局数
国立大学	101	北海道大学	69
国立短大	26	東北大学	71
国立高専	59	筑波大学	94
公立大学	91	千葉大学	51
公立短大	62	東京大学	55
公立高専	6	東京工業大学	102
大学共同利用機関	24	名古屋大学	85
国の機関	146	京都大学	75
特殊法人・独立行政法人	329	大阪大学	74
地方公共団体の機関	745	神戸大学	50
私立大学	604	広島大学	63
私立短大	517	九州大学	125
私立高専	3		
会社	4,428		
非営利団体	3,621		
その他	8		
総計	10,770		

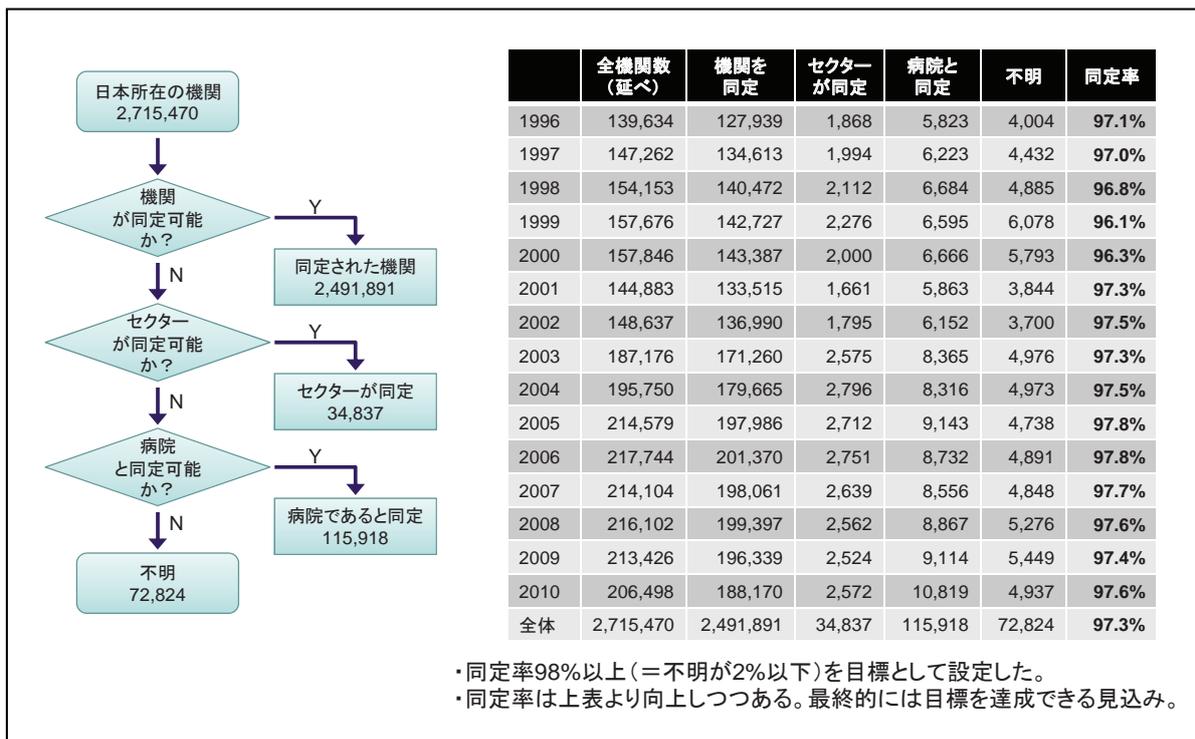
研究レビュー4-1

特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

次に、Scopus データの名寄せの状況を【資料 13】に示した。左側のフローチャートにより、名寄せの各段階の内容と結果を示している。まず、日本所在の機関を抽出することから出発する。次に、個別の機関の同定、すなわち、個別の機関としてどの機関かを特定する。このステップで9割近くが同定できている。次のステップでは、どの機関かは明確に特定できないが、どのセクターに属するか、例えば民間企業であるかどうか、などを判定する。民間企業であるかどうかは、例えば機関名称に「Co LTD」などの語句が含まれていれば判定できる。次に、病院かどうかを判定する。病院をこのように扱うのは、大学部門、政府部門、民間非営利部門、産業（民間営利）部門のいずれにも含まれることがあるため、セクターの分類とは別に、病院であることを分類しておくことが都合がよいためである。

【資料 13】 Scopus データの名寄せ結果：同定数および同定率



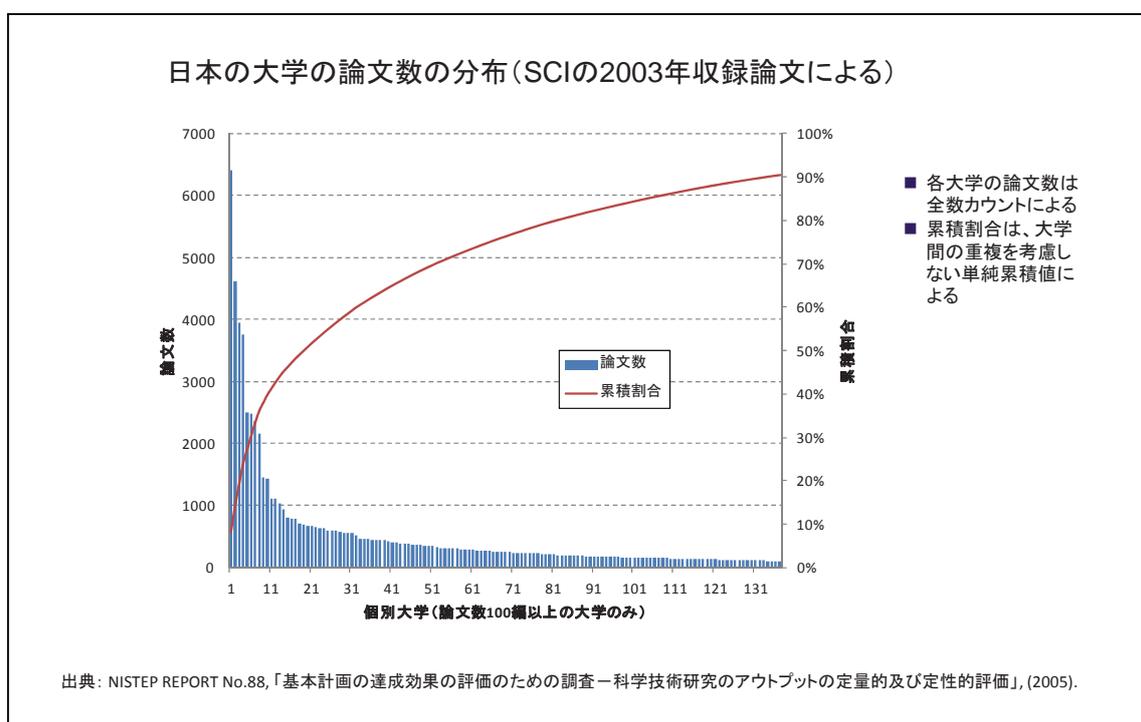
このような手順で機関の同定を行っても、不明の機関が残るが、2011年度の取り組みにおいては、不明の機関を全体の2%以下とすることを目標としていた。【資料 13】の表の作成時点では、97.3%が同定されており、不明の機関は2.7%となっている。不明の機関の割合は、今後の作業を通じて、少しずつではあるが、減らしていくことが可能であると考えられる。

1.3 名寄せされた科学論文データの分析例

前節までに述べたようなデータ整備により、どのような分析が可能になるのだろうか。現在、作成中のデータでは分析結果を示す段階には至っていないため、過去に研究として行った名寄せによるデータと分析例を示す。

第一の例は、日本の大学の論文数の分布である。個別機関別のデータの分析例としては単純なものであるが、日本の大学システムを考察する上での基礎となる情報を提供する。【資料 14】は、Web of Science (2003 年登録) から集計した個別の大学ごとの論文数を縦軸にとり、論文数の大きい順に左から並べて示している。最も論文数の多い大学では、約 6,400 編となっている。なお、この図では、2003 年に 100 編以上の論文が同定された大学のみを示しており、もし、論文数がより少ない大学についても示すとすれば、右側に長く並ぶ、いわゆるロングテールの分布となる。このような分布の形については、名寄せの精度が強く影響してくる。名寄せの精度が高くなければ、分布の図を正確に描くことはできないためである。

【資料 14】 機関別の名寄せデータの分析例：日本の大学の論文数の分布



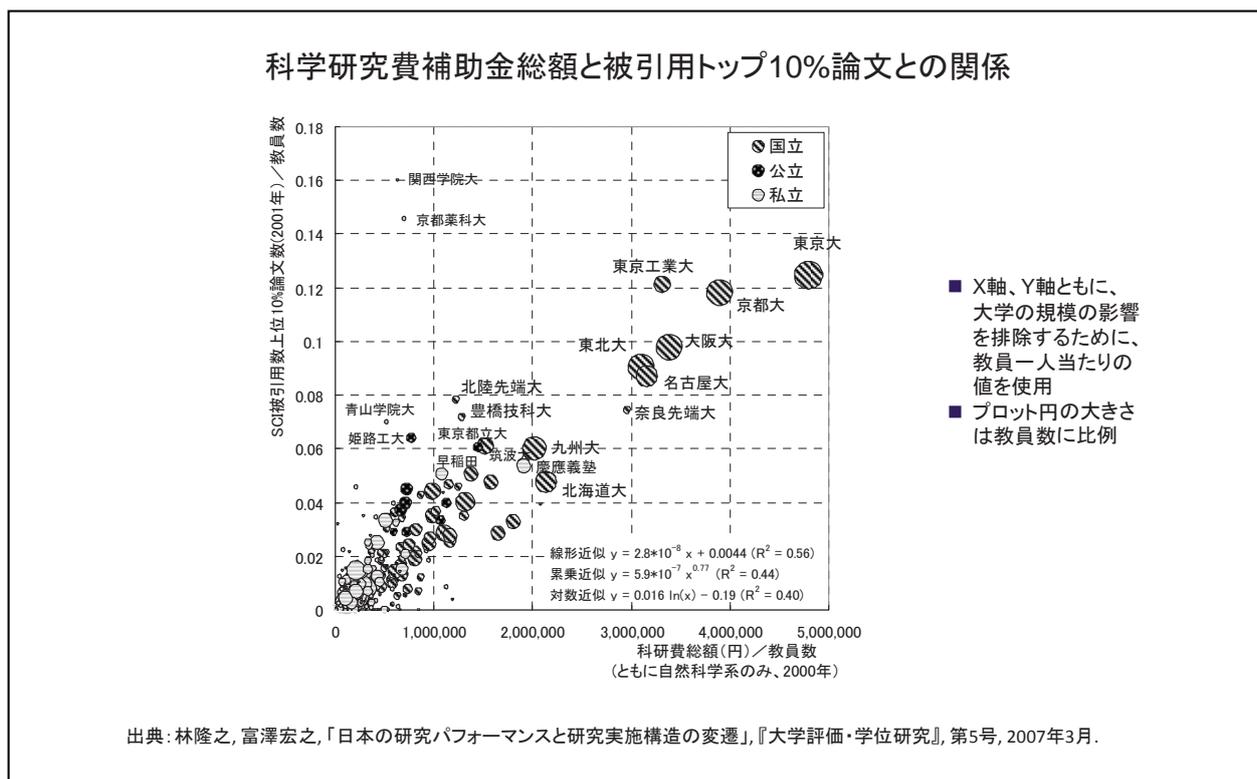
研究レビュー4-1

特許データの整合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

次に、もう少し分析を深めた例であるが、科学研究費補助金（科研費）の配分額と被引用トップ10%論文の数を、それぞれを横軸と縦軸に取り、個別大学ごとの値をプロットしたものである。ただし、科研費の総額自体とトップ10%論文数自体の間には、大学の規模による、いわゆる“見掛けの相関”が生じるが、そのような単純な関係はあまり意味がないため、縦軸、横軸ともに教員数で除して基準化している。そのような基準化をした場合でも、1人当たりの科研費の獲得額が大きい大学の方がトップ10%論文の数も多いという関係が示されている。ただし、ここに示したデータは2000～2001年についてのデータであり、現在の状況を示したものではない。

【資料15】機関別の名寄せデータの分析例：日本の大学への研究資金配分と高被引用論文数の関係

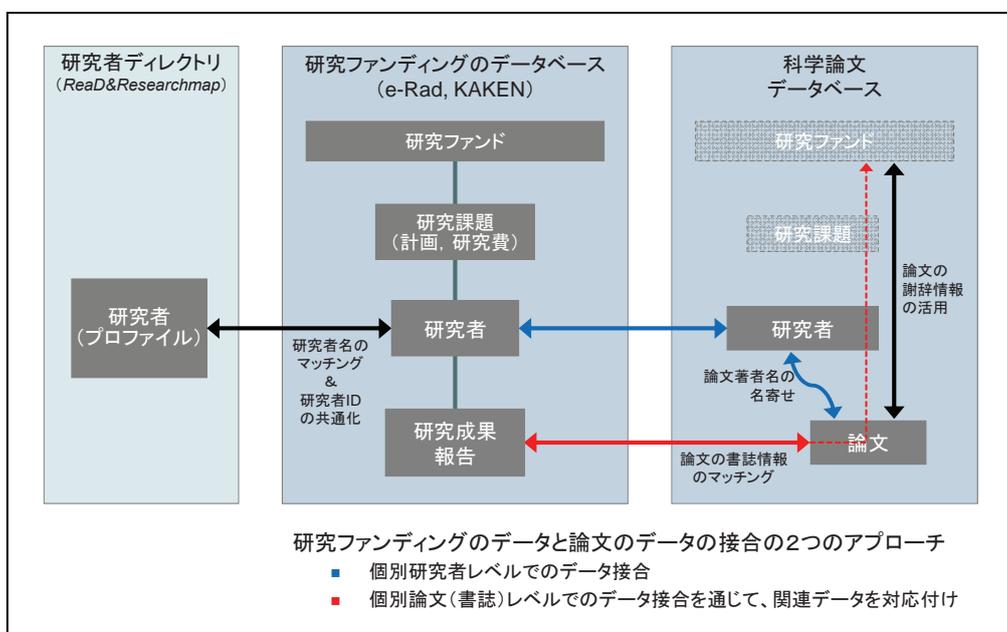


1.4 科学論文データと関連するデータの整備

ここまでに、主に科学論文データの機関レベルの名寄せについて述べたが、それ以外の関連データの整備についても言及したい。科学技術政策研究所の「データ・情報基盤の構築」において、当初より重視していたのは、政府の研究開発投資の実態や効果を把握することであった。そのため、研究ファンディングとその関連データの接合が重要であると位置付けてきた。

【資料 16】は、その概念モデルを示している。図の中央に、研究ファンディングのデータベースが配置されている。これに該当するものとしては、日本政府の府省共通の「e-Rad」という研究ファンド等の管理システムや、科研費に関するデータを収録した「KAKEN」という良く整備されたデータベースもある。このような研究ファンディングのデータベースには、研究ファンドの種類ごとに、研究課題や配分された研究費、あるいは研究者についての情報、さらに、研究ファンドにより実施された研究成果（論文など）についての情報などが収録される。ただし、研究成果の情報の収録状況はファンドによって異なる上に、個別の研究課題の成果把握が目的で収集されることが一般的であるため、論文の書誌情報の書き方が統一化されていなかったり、同一の論文の情報が重複して収録されるなどの問題があり、研究ファンドの全体的なアウトプットについての体系的な分析に適していない。

【資料 16】 機関レベル以外のマイクロレベルのデータ接合：研究ファンディングとその関連データ



研究レビュー4-1

特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

そのような体系的な分析のためには、科学論文のデータベースを用いることが望ましいため、研究ファンドの成果リストを論文データベースと接続させることが重要になる。この点について、【資料 16】の下方では、「研究成果報告」と「論文」の間をつなぐ赤い矢印により、個々の論文レベルでの両者の連結を表現している。このような論文書誌レベルでデータを接合させるためには、研究成果報告に記載された論文の書誌情報を、論文データベースに含まれる書誌情報と照合させ、同一の論文のデータを見つけ出す必要がある。それをコンピューターで自動的に行うためのマッチングのアルゴリズムについては、著者を含む何人かの研究者が開発を進めている。

別のレベルでのデータ接合としては、「研究者」のレベルでのデータ接合が考えられる（図では、青い線の矢印で示している）。そのための一つの方法は、論文のデータを著者名で名寄せすることである。図中の「科学論文データベース」のなかの「論文」と「研究者」をつなぐ青い矢印がそれを表している。このように研究者ごとの論文データを作成すれば、研究ファンドのデータベースに収録されている研究者ごとのデータと接合することが可能となる。例えば、科研費の場合、科研費を獲得した研究者ごとに、それぞれの研究内容や研究費金額に加えて、成果論文を把握することができる。

研究者のレベルでのデータ接合には、もう一つのアプローチが考えられる。研究ファンドのデータベースには、研究者と研究成果（論文など）の結びつきの情報が収録されているので、それを科学論文のデータベース側に移入すれば、科学論文の著者についての名寄せを行う必要は無い、という考え方である。ただし、このようなことが可能になるためには、研究ファンドのデータベースにおいて、研究成果が研究者ごとに整理されていることや、研究者に ID が付与されているなどの一意性が確保されている必要がある。実際には、そのようにデータが収集されているとは限らない。また、研究ファンドを受けずに実施される研究（大学の基盤的経費による研究など）との比較ができない、などの制約がある。

なお、1.1 節で科学論文データベースの実例を示した際に、データベースの謝辞情報のなかに、研究ファンドや獲得した研究費の情報が含まれていることを述べた。この図では、科学論文データベースの一番下の階層に「論文」があり、それが最上位階層の「研究ファンド」と縦の黒い線の矢印でつないでいる。このように、データベース中の謝辞情報を活用して、研究ファンディングについて把握しようとするアプローチについても検討している。ただし、謝辞情報はデータベースに常に収録されているわけではなく、もともと、論文において謝辞の書き方にばらつきがあるため、現時点では、質の高いデータではない。

以上に述べたようなデータ接合における方法論上のチャレンジについても、簡単に触れておく【資料 17】。論文著者の名寄せは、先に述べた機関の名寄せに比して困難度が高い。そもそも、人名のみでは個人を完全に同定することはできない上に、英語論文の場合、日本人名の漢字に含まれる情報の欠落といった問題がある。そのため、論文著者の名寄せに際しては、人名

特許データの統合的データ体系の構築：データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

に加えて、所属機関や研究分野の情報を用いることが必要になってくる。その場合でも、所属組織の異動や同一組織内での同姓同名があれば、同定は著しく困難である。そこで考えられる一つのアプローチとしては、共著者の類似性を判断材料にする方法が考えられる。すなわち、同じ人名であるが異なる機関に所属する2人が同一人物かどうかを判定するために、どのような共著者と一緒に論文を書いたかを参照するという方法である。科学技術政策研究所の「データ・情報基盤の構築」では、以上のような方法により、論文著者の名寄せを実施している。ただし、これは試行的な取り組みであり、今後、一層の精度向上が必要である。

【資料 17】 機関レベル以外のマイクロレベルのデータ接合：データ接合のためのチャレンジ

- 論文著者の名寄せ (= 研究者の同定)
 - 人名のみでは、完全な同定は不可能
 - 同姓同名の問題
 - 英語論文の場合、日本人の漢字情報が欠落
 - 論文データベースの古いレコードでは、ファーストネームがイニシャルのみ
 - 所属機関、研究分野、E-mailなどの情報の活用は不可欠
 - 所属の変更(異動)の場合、同定が困難
 - 共著者の類似性による同定は、ある程度、有効
 - NISTEPの2011年度のデータ整備では、上記の全てを活用した名寄せを実施
- 論文(書誌情報)のマッチング
 - 論文の(1)ジャーナル名、(2)巻号、(3)ページ、(4)出版年、がマッチング・キーとしてよく用いられる。(5)著者名、(6)論文のタイトル、も用いられる場合がある
 - (6)については、テキストの類似度を用いる必要がある
 - NISTEPが開発したマッチング・アルゴリズムは(1)~(6)の様々な組み合わせ
 - 研究ファンディングの成果報告書中の論文の同定、特許に引用された科学論文の同定などに有効

もう一つの方法論上のチャレンジは、既に言及した論文のマッチングである。これは、前述の研究ファンドの成果リストと論文データベースとの接続が代表的な例であるが、異種のデータベースにおいて、同一の論文のデータが含まれているかどうかをコンピュータ・プログラムにより同定する方法である。通常、このような論文書誌情報のマッチングでは、論文のジャーナル名、巻号、ページ、出版年を用いる。しかし、著者は、それに加えて著者名や論文のタイトルも用いる方法の開発を進めている。特に、論文のタイトルについては、テキストの類似度を用いたマッチング方法の開発を進めている。この方法が確立できれば、前述の研究ファンディングの成果報告書のなかの論文の同定に加えて、特許に引用された科学論文の同定といったことも可能になる。

研究レビュー4-1

特許データの統合的データ体系の構築:データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

以上、第1節では、科学論文データの名寄せを中心に述べてきた。【資料18】は、2012年度のデータ整備の構想を示している。これは現時点での構想であり、最終的な決定内容を示したものではないが、基本的な方針としては、2011年度に開始した機関名辞書の作成などを継続し、データの精度向上と範囲拡大に向上に取り組む。特に、科学論文のデータベースについては、2011年度はScopusのみを対象としたが、2012年度はWeb of Scienceも対象とすることを検討している。また、日本だけでなく海外の機関についても名寄せを行うことも検討している。

【資料18】大学・公的研究機関に関するデータ整備

- 2011年度に整備した機関名辞書などのデータ提供を開始予定
 - 2012年夏頃に最初のデータ提供を開始する予定
- データの向上(精度向上・範囲拡大)に継続的に取り組む
 - データのメンテナンス(データ更新など)にも取り組む
- 科学論文DBとしてWeb of Scienceも対象にすることを検討中(2011年度はScopusのみを対象)
- 日本だけでなく、海外主要国の論文データの機関名の名寄せを行うことを検討中
 - ある程度網羅的な大学のリストが入手できる国を対象にする
 - 具体的には、米国、イギリス、ドイツ、中国などが候補
- 論文著者の名寄せ
 - 2011年度に試行した論文著者の名寄せの精度向上に取り組む予定
- e-RadやKAKENとReaD&Researchmapの連携の推進

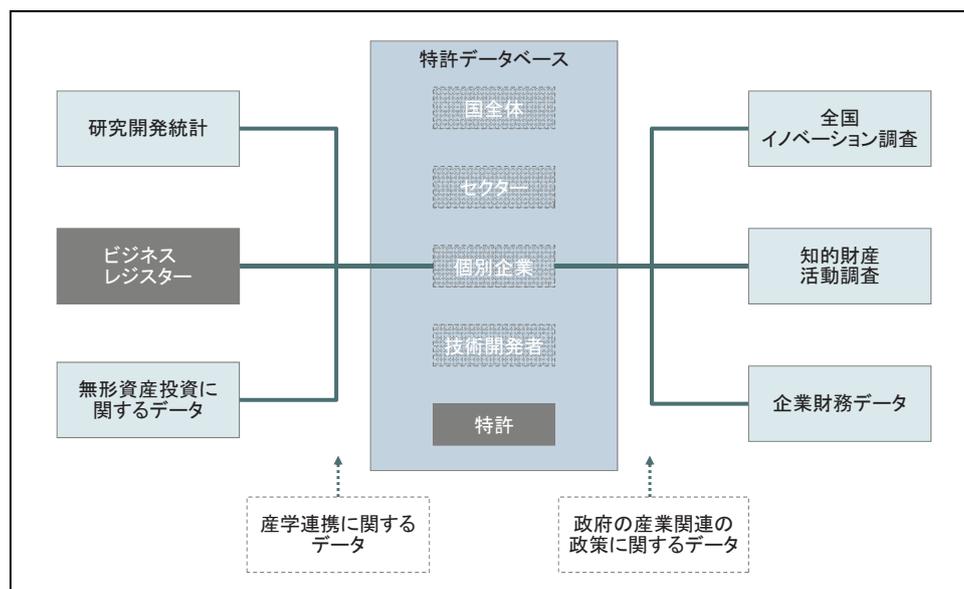
特許データの整合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

2. 特許データにおける名寄せ

データの名寄せが重要となるのは科学論文のデータに限らない。以下では、特許やその関連データに関するデータ整備についても簡単に触れておきたい。特許を中心としたデータ整備の狙いは、産業における研究開発や技術的イノベーションについての理解を深めることである。特に、産業における研究開発やイノベーションに関して、各種の統計をはじめとする様々なデータを個別企業レベルで連結するとともに、そのなかに特許データを主要データとして位置づけようとするものである。このような各種データの個別企業レベルでの連結については、政府統計の中でビジネスレジスターと呼ばれる網羅的な企業リストの整備の計画があり、今後、それが実際に機能するようになると、各種データの個別企業レベルでの連結が、特段の困難なく実現する可能性もある。しかし、そのような事が実現するのは、しばらく先のことになると考えられ、また、過去に遡ってデータを分析するためには、企業名についての名寄せが必要である。

【資料 19】 産業における研究開発・イノベーションに関するデータ整備の概念モデル



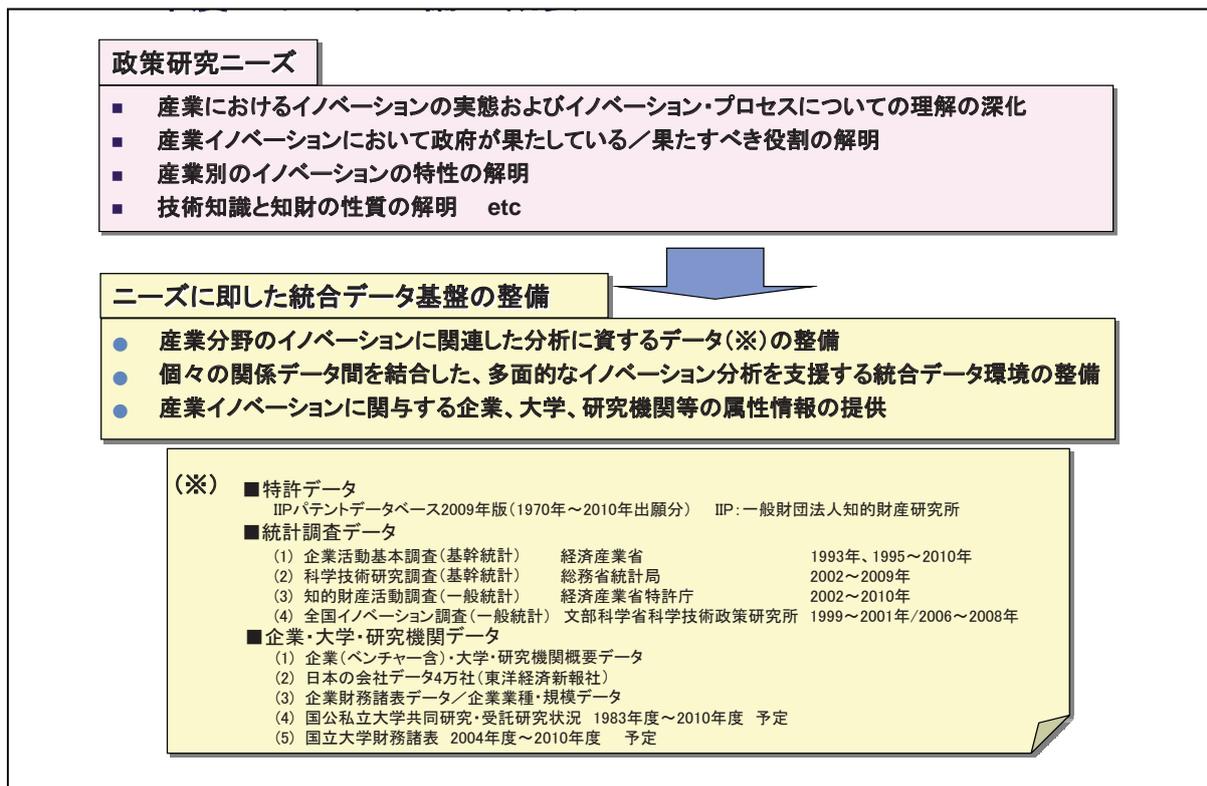
特許データの統合的データ体系の構築: データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

以下の【資料 20】と【資料 21】には、NISTEP が取り組んでいる「産業における研究開発に関するデータ整備」の概要を示している。簡単に言えば、特許データを中心に企業の名寄せを行い、企業名辞書を中心に置き、それとさまざまなデータが相互につながるようにしようとするものである。

なお、2012 年度のデータ整備に関しては、まず 1 年目に整備したデータが非常に複雑であるため、データの精査が必要であり、また、試行的に分析することにより、データ整備の仕方について、評価する必要がある。そのようなデータの精査等の後、機関名辞書などのデータ公開を開始する予定である。また、その後もデータのメンテナンスには継続的に取り組む予定である。

【資料 20】 産業における研究開発に関するデータ整備：2011 年度のデータ整備の概要



特許データの整合的データ体系の構築：データの名寄せの挑戦

科学技基盤調査研究室 富澤 宏之

【資料 21】 産業における研究開発に関するデータ整備：2011 年度のデータ整備の概要

■統合データ基盤

企業の事業活動に必要な研究開発活動、知的財産化活動の状況などに関し、個別の目的で調査・整備されてきた各種の統計データや特許データ等を産業イノベーションという視点で結合し、産業の創造と発展に向けた政策研究ニーズに応える新たな価値ある情報を生み出すプラットフォームとして活用可能とする

■結合の鍵とする情報

統計調査は企業単位の実施、特許データは特許出願単位のデータセットであることから、共通のキーとして調査企業名／出願企業名を名寄せし結合する

■企業名の名寄せ

- 蓄積された統計・特許データに記載される企業名には、法人格表記・略称・新旧漢字等に起因する表記揺れ、誤記述、過去の社名、同名異企業など名寄せの障害要因を数多く抱えている
- このため、名寄せの正確性の向上を担保すべく、企業名データのクレンジング、法人格表記の標準化を行い情報の品質を高め、False Negative(同一企業を異企業とみなしてしまう)、および、False Positive(異なる企業を同一企業と見なしてしまう)の問題の最小化に向けて、住所情報も付加した照合を行う

例 IIPパテントデータベースの企業出願人の名寄せ推移

出願人企業レコード総数:	303,483
企業名クレンジング・正規化:	157,796
企業名+所在地12桁コード:	185,226

3. データ整備がもたらすもの: 増大するデータの活用の試み

最後に、データの名寄せというテーマから離れるが、本稿で述べたようなデータの整備の意義について、より広い視野からの考察を加えておきたい。

世界的な“情報爆発”が起きていると言われることがあるように、現代社会において、データは増大する一方である。そのような増大するデータを本格的に活用しようという動きが表れている。例えば、米国では、「Data to knowledge」や「Data to decisions」というコンセプトのプロジェクトが開始されている。前者はNSF、後者は国防省のプログラムのキーワードであるが、この語からうかがえるように、データを固有の知識の源であると捉えたり、意思決定のために活用するといったコンセプトの新しいプログラムである。科学技術政策研究所の「データ・情報基盤の構築」は、文部科学省の科学技術イノベーション政策における「政策のための科学」推進事業の一環であるが、このような海外における最新の取り組みを見ると、当初、考えていた以上に広がりのある試みになる可能性もあるのではないだろうか。

【資料 22】 増大するデータ、知のためのデータ、意思決定のためのデータ

- 世界的な“情報爆発”
 - 社会の情報化が大量の情報を産み出しており、その量は急激に増加
 - インターネットを通じて世界的に流通
- 増大するデータを本格的に活用しようとする動き
 - “Data to knowledge” [“From data to knowledge”] (NSF, U.S.)
 - “Data to decisions” (DoD, U.S.)
- 「政策のための科学」の根幹としてのデータ整備
 - 科学技術イノベーション活動の測定は本質的に困難であり、これまで限定的にしか、行われていなかった…
 - 増大するデータ・情報の活用によって、ようやく本格的な定量的把握が可能に？
 - 意志決定の質を決定付ける要因としてのデータ活用



富澤 宏之
科学技術基盤調査研究室長

(経歴)

1988年 日本科学技術情報センター
1989年～1996年 科学技術庁科学技術政策研究所研究員
1996年～2006年 同研究所 主任研究官
2006年～ 同研究所 科学技術基盤調査研究室長
(現在に至る)

(この間)

1992年～1993年 東京大学大学院総合文化研究科研究員
1995年 欧州委員会未来技術研究所 (IPITS) 滞在研究員
2007年～2010年 経済協力開発機構 (OECD) 科学技術産業局
(STI) 主席行政官

参考文献:

[1] NISTEP NOTE (政策のための科学) No. 3

「科学技術イノベーション政策のための科学」におけるデータ・情報基盤構築の推進に関する検討 (2012年11月)

<http://hdl.handle.net/11035/1176>

[2] 富澤宏之, 「科学技術イノベーション政策に有用なデータ基盤は何か: 中長期的構想」, 研究・技術計画学会 2011年次学術大会講演要旨集, p.86-89 (2011年11月)

[3] 元橋一之, 鈴木潤, 「エビデンスベースの科学技術政策の実現に向けたイノベーションデータベースの構築」, 研究・技術計画学会 2010年次学術大会講演要旨集, pp. 225-228 (2010年10月)

[4] 元橋一之, 鈴木潤, 「イノベーションデータベースの開発と研究者に対する提供」, 研究・技術計画学会 2008年次学術大会講演要旨集, pp. 522-525 (2008年10月)

[5] 調麻佐志, 藤垣裕子, 平川秀幸, 富澤宏之, 林隆之, 牧野淳一郎, 『研究評価・科学論のための科学計量学入門』, 丸善株式会社, 2004年.

[6] 林隆之, 富澤宏之, 「日本の研究パフォーマンスと研究実施構造の変遷」, 『大学評価・学位研究』, 第5号, 2007年3月, pp.3-19. (塚原修一編著『高等教育』, 日本図書センター, 2009年9月に再録)

[7] Hiroyuki TOMIZAWA and Takayuki HAYASHI

“Constructing a Multi-level Scientometric Indicators System,” Blue Sky II 2006 Forum: What Indicators for Science, Technology and Innovation Policies in the 21st Century? September 25–27, 2006, Ottawa, Canada.

[8] NISTEP REPORT No.88

基本計画の達成効果の評価のための調査—科学技術研究のアウトプットの定量的及び定性的評価—（平成 15 年度～16 年度科学技術振興調整費調査研究報告書）

<http://hdl.handle.net/11035/623>

〔講演〕

特許情報を使った科学技術イノベーション
政策研究の事例報告

一橋大学 長岡 貞男

講演

特許情報を使った科学技術イノベーション政策研究の事例報告

一橋大学 長岡 貞男

はじめに: 科学技術イノベーション政策の分析における特許情報の重要性

科学技術イノベーション政策の目標として以下が重要だと考えられる。

- サイエンスとイノベーションのシナジーを実現
- イノベーションへの科学的な源泉の充実(“science capital for innovation”)
- サイエンスを刺激するイノベーションの推進(Pasteur quadrant)
- 科学を基盤とする新産業分野の開拓ができる人材や企業の育成

こうした目標に照らして、現実の動きを評価する上で、特許情報は重要である。

その理由:

- (1)サイエンスの成果をイノベーションに生かしていく過程の中で、中間的な成果が発明。

Science → invention → innovation

- (2)このため、特許情報は、以下を把握する上で重要な情報を提供。

- 科学を活用した発明の創造過程、
- 発明から新製品・新商品のイノベーション創出過程

講演

特許情報を使った科学技術イノベーション政策研究の事例報告

一橋大学 長岡 貞男

1. 特許情報の基本的内容

特許情報に何が含まれているかを理解する上で、先ず、発明と特許権とは何かを理解しておく必要がある。

発明(invention): 経済的に有用な新たな技術的知識

「特許法 第二条 この法律で「発明」とは、自然法則を利用した技術的思想の創作のうち高度のものをいう」(特許法)

特許権: 発明をその請求項の範囲で排他的に実施する権利

特許性の要件: 新規性、進歩性、有用性

特許情報に含まれる主要な情報は以下の通りである。具体的な例として、京都大学の山中伸弥先生の iPS 細胞の製造方法の特許公開公報の書誌事項を【資料 1】に示している。これらの情報は特許公開公報及び特許登録公報によって一般に開示されている。

- 発明
- 発明者
- 出願人
- 優先日、出願日、公報日
- 特許分類
- 請求項
- 発明の記述
- 特許庁

【資料1】 i P S細胞の製造法の特許公開公報

(19)【発行国】日本国特許庁(JP)
(12)【公報種別】公開特許公報(A)
(11)【公開番号】特開2008-283972(P2008-283972A)
(43)【公開日】平成20年11月27日(2008.11.27)
(54)【発明の名称】誘導多能性幹細胞の製造方法
(51)【国際特許分類】

C12N 5/10 (2006.01)
C12N 15/09 (2006.01)
A61K 35/12 (2006.01)
A61L 27/00 (2006.01)
A61P 25/16 (2006.01)
A61P 3/10 (2006.01)
A61P 35/00 (2006.01)

【F I】

C12N 5/00 ZNA B
C12N 15/00 A
A61K 35/12
A61L 27/00 V
A61P 25/16
A61P 3/10
A61P 35/00

【審査請求】有
【請求項の数】1
【出願形態】OL
【全頁数】44
(21)【出願番号】特願2008-131577(P2008-131577)
(22)【出願日】平成20年5月20日(2008.5.20)
(62)【分割の表示】特願2007-550210(P2007-550210)の分割
【原出願日】平成18年12月6日(2006.12.6)
(11)【特許番号】特許第4183742号(P4183742)
(45)【特許公報発行日】平成20年11月19日(2008.11.19)
(31)【優先権主張番号】特願2005-359537(P2005-359537)
(32)【優先日】平成17年12月13日(2005.12.13)
(33)【優先権主張国】日本国(JP)
(71)【出願人】
【識別番号】504132272
【氏名又は名称】国立大学法人京都大学
【住所又は居所】京都府京都市左京区吉田本町3-6番地1
(74)【代理人】
【識別番号】110000109
【氏名又は名称】特許業務法人特許事務所サイクス
(72)【発明者】

講演

特許情報を使った科学技術イノベーション政策研究の事例報告

一橋大学 長岡 貞男

【氏名】 山中伸弥
【住所又は居所】 京都府京都市左京区聖護院川原町5-3 国立大学法人京都大学再生医科学研究所内
【テーマコード（参考）】
4B024
4B065
4C081
4C087
【Fターム（参考）】
4B024 AA01 BA80 CA01 DA02 EA02 GA11 GA18 GA23 GA30 HA17
4B065 AA90X AA90Y AB01 AC12 BA02 BA06 CA44
4C081 AB11 CD34
4C087 AA02 AA04 BB33 BB57 NA14 ZA02 ZB26 ZC35

(57) 【要約】 (修正有)
【課題】 胚やES細胞を利用せずに分化細胞の初期化を誘導し、ES細胞と同様な多能性や増殖能を有する誘導多能性幹細胞を簡便かつ再現性よく製造する方法を提供する。
【解決手段】 体細胞から誘導多能性幹細胞を製造する方法であって、下記の4種の遺伝子：Oct3/4、Klf4、c-Myc、及びSox2を体細胞に導入する工程を含む方法。
【選択図】 図1

【特許請求の範囲】
【請求項1】
体細胞から誘導多能性幹細胞を製造する方法であって、下記の4種の遺伝子：Oct3/4、Klf4、c-Myc、及びSox2を体細胞に導入する工程を含む方法。

また、特許権の権利についての以下の情報も公表されており、これらは特許発明がどの程度「利用」されているか(そもそも審査請求をされたのか、どの程度の期間、権利を維持されたのか)、また特許権の移転が行われたか、またそもそも特許権がどの程度安定しているのか(無効審判等で無効になっているかどうか)等について重要な情報を提供している。

- 権利の成立、移転、消滅にかかる情報
- 権利の成立過程 審査請求、査定と登録、無効・不服審判(異議申し立て)
- 権利の移転 権利の譲渡や合併・買収に伴う権利の移転
- 権利の維持期間

2. 発明の企業秘密による保護と特許出願

発明は特許出願されるとは限らない。企業には以下の選択肢がある。

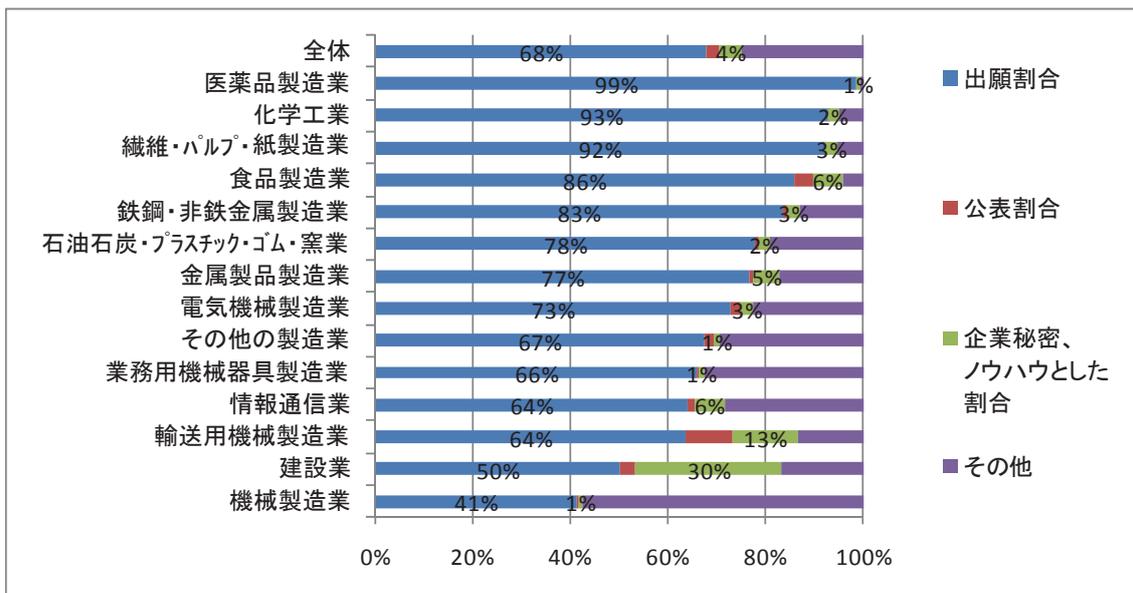
- 特許出願
- 公表のみ
- 企業秘密・ノウハウとしての保護

企業秘密による保護が選択されている発明がどの程度存在するかが、重要であるが、先ず発明数自体を把握する包括的な統計は存在しない。原理的には、以下の三つに分けることができる。

- 知的財産部に届けられた発明
- 事業部が把握している発明
- 発明者が把握している発明

以下の【資料 2】は、知的財産部に届け出のあった発明の内、どの程度が出願され、どの程度が公表され、どの程度が企業秘密となっているかを示す。平均では、発明の中で約 7 割が出願され、企業秘密・ノウハウとしたケースは 4% である。

【資料 2】 知的財産部に届け出のあった発明の出願割合と企業秘密・ノウハウとした割合



出典： 知的財産活動調査、特許庁、2008 年

科学的な貢献が大きい重要な発明は特許となっている可能性が高いと考えられる。その理由として以下を指摘できる。

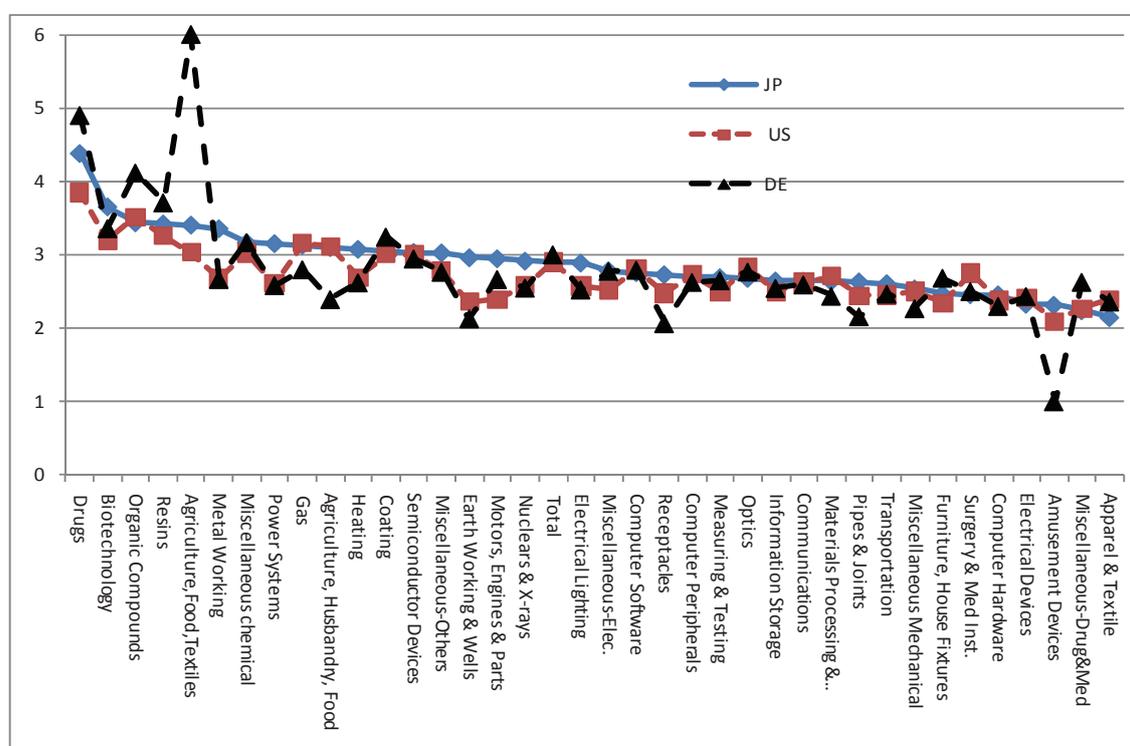
- 研究開発競争の存在
- 特許出願をしないと、他社によって特許化され、自らの発明の利用ができなくなる危険性がある。後者が重要だと各企業によって特許出願がドミナント戦略となる。

3. 発明者と発明チーム

特許情報の中で最も重要な情報の一つは発明者についての情報である。これは例えば、発明におけるチームの重要性が高まっているかどうか等の問いを分析するのに利用できる。特許当たりの発明者数は徐々に増加している(Jones, B. F. (2009)は”Burden of knowledge?”と指摘している)。また、以下の【資料3】に示すように、サイエンスとの関連性が強い分野で平均発明者数が多い。

こうした発明者の情報を利用して、産学連携を含めて、組織を超えた協力が発明においてどの程度広範に行われているかどうか、また国際的な人材や知識を有効に使っているかどうか等を分析することが可能である。

【資料3】 技術分野別の平均発明者数(三極出願)



出典： Tsukada Naotoshi and Sadao Nagaoka(2012)

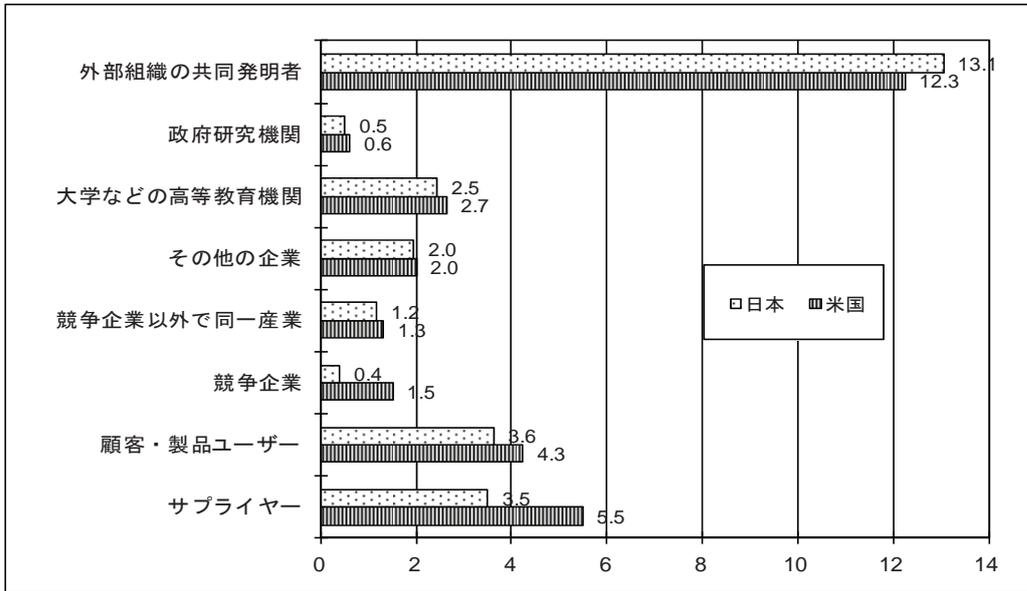
特許情報のみから発明者の所属組織の関係を知ることは必ずしも容易ではない。発明者にサーベイをすることで、これを知ることができるが、日米とも、【資料4】によると三極出願特許¹について1割強の発明には外部組織の共同発明者が存在していることが知られている。

組織間の協力の類型をみると、垂直的な関係(ユーザー、サプライヤーによる共同発明)が日米とも最も頻度が高い。また、大学、国立研究機関と産業界との共同発明は、日米ほぼ同水準であり、水平的な関係(競争企業)の共同発明は日米とも頻度が低い。

¹ 三極出願特許は米国で登録され、日本国特許庁及び欧州特許庁に出願された特許。

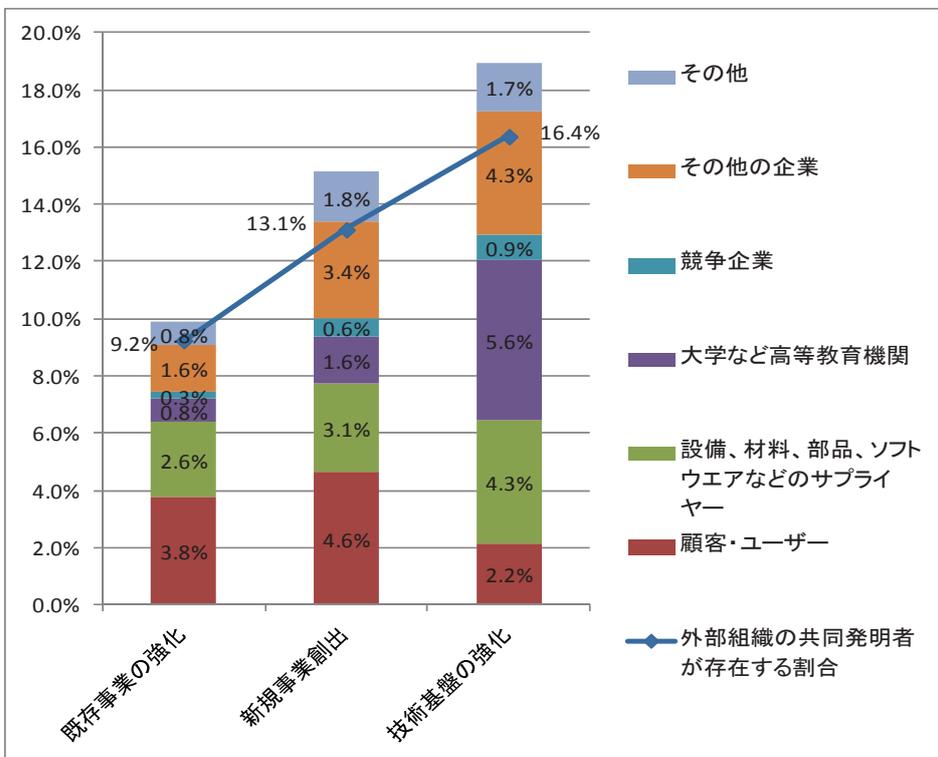
【資料 5】が示すように、企業にとって現在の事業との距離が大きい分野での研究開発において、協力関係はより広がる。

【資料 4】 外部組織からの共同発明者の存在頻度(日米からの三極出願特許、%、共通の技術分野構成)



出典: Walsh and Nagaoka (2009a)

【資料 5】 研究開発の目的と共同発明者の所属組織



出典: 長岡(2011)

講演

特許情報を使った科学技術イノベーション政策研究の事例報告

一橋大学 長岡 貞男

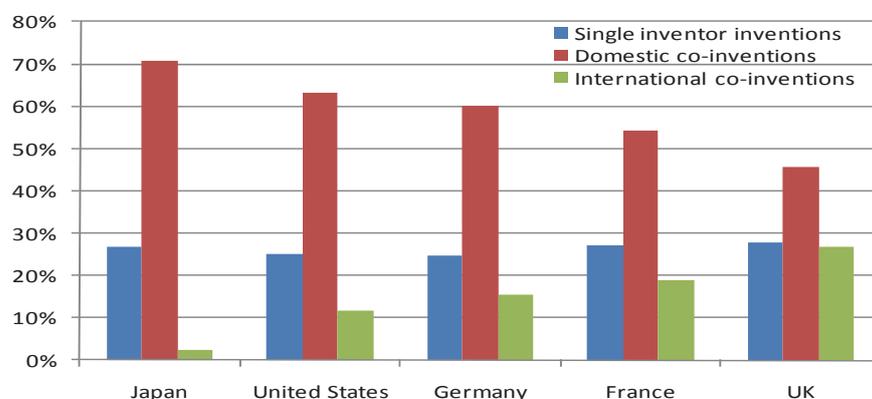
4. 国際共同発明

グローバルに知識を求めることが重要となっており、国際共同発明も重要となっている。特許情報を使ってこれを把握することも可能である。それを活用した分析によれば、【資料6】が示すように、日本の発明(発明者および出願人それぞれに日本人あるいは日本法人が存在)には、諸外国と比較した国際共同発明が少ない。以下の二つの潜在的な要因が指摘できる。

- 各企業内において国際的な人材を利用していない
- 外国企業との国際共同研究が少ない

【資料7】に示す統計的な分析の結果によれば、前者の方がより重要である。

【資料6】 国際共同発明の割合



出典: OECD の Triadic patents (2000-2006)

【資料7】 発明者と出願人の構造(2000-2006)、三極出願特許

Japan (2000-2006)

	Single inventor	Domestic co-inventions	International co-inventions	Total
Single applicant	25.9%	64.1%	1.0%	91.0%
Domestic co-applications	0.8%	7.3%	0.1%	8.2%
International co-applications	0.1%	0.2%	0.5%	0.8%
Total	26.8%	71.7%	1.5%	100.0%

United States (2000-2006)

	Single inventor	Domestic co-inventions	International co-inventions	Total
Single applicant	25.0%	62.7%	7.0%	94.7%
Domestic co-applications	0.2%	3.3%	0.2%	3.7%
International co-applications	0.1%	0.4%	1.2%	1.6%
Total	25.3%	66.4%	8.3%	100.0%

Germany (2000-2006)

	Single inventor	Domestic co-inventions	International co-inventions	Total
Single applicant	25.2%	61.3%	7.6%	94.1%
Domestic co-applications	0.4%	3.5%	0.2%	4.1%
International co-applications	0.1%	0.3%	1.4%	1.8%
Total	25.6%	65.2%	9.2%	100.0%

出典: Nagaoka and Tsukada(2012)

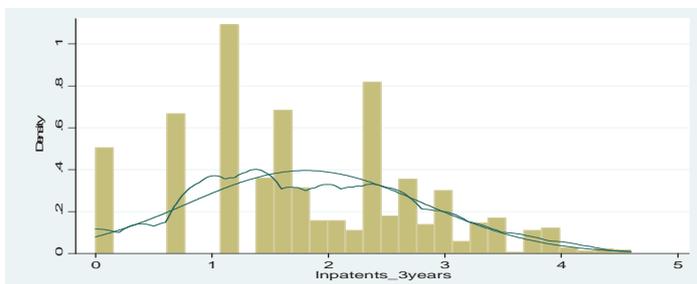
5. 発明者の生産性

発明の生産性を高めるための教育や制度の在り方の検討にも、特許のデータは重要である。比較的少数の発明者に発明が集中していることはよく知られている。Narin, and Breitzman(1995)によれば、“the key role of a few researchers seems to be a law of nature.”である。日本の企業の発明者の3年間の特許出願数の分布を示したのが【資料8】である。1件から50件以上と幅が非常に大きく、ほぼ対数正規分布に従っている。3年ではなく、ライフタイムでも同様な傾向がある。

このような大きな格差の原因として、発明者が発掘した「鉱脈」の大小、良い成果がある発明者に優先的に資源が割り当てられるための累積的な優位性、能力などが考えられるが、その原因はまだ明確ではなく、今後その分析も重要である。

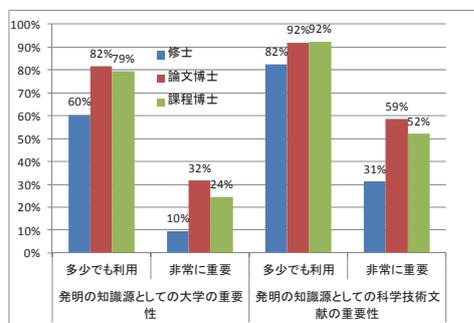
また、高等教育、組織間の移動、人事制度のあり方なども重要である。サイエンスを活用する能力を企業の発明者が持つことが重要であり、そのためには自らも科学的な研究成果ももたらすことができる博士レベルの人材が重要である。【資料9】は、博士(課程博士、論文博士)の発明者が大学(大学の研究者など)と科学技術文献を、それぞれ発明の知識源として活用した程度を修士卒の発明者と比較して示している。論文博士と課程博士の間で、サイエンスの成果を吸収する能力に大きな差はないことが注目される。

【資料8】 1人当たりの特許出願数の分布、(3年間、対数、カーネル分布と正規分布近似)



出典：長岡、塚田、西村、大西(2012)

【資料9】 大学や科学技術文献を発明の知識源として活用する程度 (修士、課程博士および論文博士)



注「多少でも利用」は、利用しなかった場合以外の場合すべてである。

出典：長岡、塚田、西村、大西(2012)

講演

特許情報を使った科学技術イノベーション政策研究の事例報告

一橋大学 長岡 貞男

6. 知識ストックと知識スピルオーバー

知識の特性は、利用における非競合性があることである：

- 多数の企業が同時に利用することが出来る
- 消費されず、一度生産された知識は永久に利用可能

また、知識の生産に最も重要なインプットは知識自体である。

更に、特許の利用に関して、試験研究の例外があり、当該発明を進歩させるために発明を利用することは特許侵害にはならない。科学技術論文の成果は自由に利用することができる。したがって、ある経済主体によって開発された知識が、対価を伴うことなく他の経済主体によっても利用することが可能であり、知識のスピルオーバーが存在する。

このことが、知識ストック=過去に全ての主体で創造された有用な知識の全体の概念が、経済成長において非常に重要な原因である。

知識ストック 一部の知識は新しい知識の創造によって陳腐化しても、知識ストックは常に増加。

- 知識ストックの拡大が持続的な経済成長の源泉
- 知識ストックの存在によって技術開発の一定の方向性

政策課題として、知識ストックの範囲と質、その利用の水準を高めることが経済成長との関係で重要である。

- 開示の質
- 海外の知識がどの程度利用されているのか
- 吸収能力・連携能力

特許の査定は発明の開示が条件であるが、以下の【資料10】に見るように、技術分野によって、その有用性に大きな差がある。ソフトウェアの分野では、研究の着想源として、科学技術文献は重要であるが、特許文献は重要ではないとの結果になっており、開示制度に問題があるように考えられる。

【資料10】 研究開発の着想における特許文献と科学技術文献の重要性(「非常に重要」の割合、%)

	日本	非常に重要な割合(%)		米国	非常に重要な割合(%)	
		特許文献	科学技術文献		特許文献	科学技術文献
最も重要	樹脂	37	19	医薬品	33	51
	医薬品	32	51	樹脂	32	26
	有機化合物	32	31	手術・医療機器	24	26
	コーティング	31	27	その他の化学	23	23
	バイオテクノロジー	30	47	コーティング	19	54
	全てのセクターの平均	23	19	全てのセクターの平均	15	20
最も重要でない	測定・試験	17	30	その他の電気機器	7	10
	光学	16	11	情報ストレージ	5	18
	半導体部品	16	21	モーター、エンジン、同部品	4	10
	情報ストレージ	16	16	金属加工	2	7
	コンピューター・ソフトウェア	10	16	コンピューター・ソフトウェア	1	14

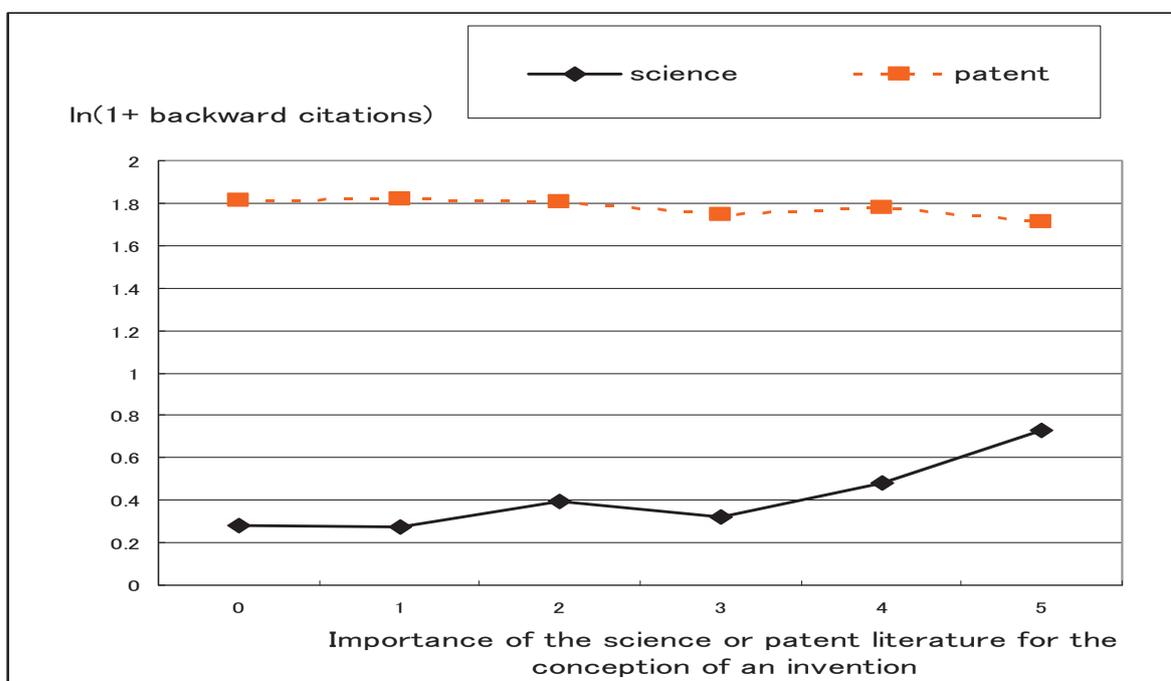
日米発明者サーベイから作成

特許情報(引用情報)を利用して、サイエンスからの知識フローの評価が活発に行われているが、誰が何故、引用しているのか、国際的に引用のメカニズムは異なっており、それを理解することが重要である。登録特許の参考文献は当該発明者の研究の着想に重要な情報かではなく、当該発明の当該専門分野全体から見て当該発明の新規性・進歩性の評価に適切な文献が掲載されている。またこうした文献を識別する作業は、日本、欧州では主として審査官が行なっており、米国では、審査官と出願人の両方が行なっている(Thompson (2006))。また、出願人が引用文献を選定する場合も、発明者ではなく、知的財産部や弁理士である場合もある。

こうした結果、米国における発明者のサーベイによる研究結果(Jaffe, Adam B., Manuel Trajtenberg, and Michael S. Fogarty(2002))が示すように、発明者が事前に知らない特許引用文献が半分以上であり、知っている場合も、特許出願の段階となって知った場合が多い。したがって引用特許文献の多くは、発明者への直接的な知識フローを示さない。

以下の【資料 11】が示すように、米国特許において特許文献の引用と科学技術文献の引用では、知識フローの指標としての重要性は後者の方が高い(後者の方が知識源としての重要性と相関が高い)。審査官が科学技術文献を引用する頻度は比較的に低いことを反映していると考えられる。

【資料 11】 研究開発プロジェクトの着想における特許文献と科学技術文献の重要性とそのプロジェクトからの特許が引用している特許文献と科学技術文献の数



出典：Nagaoka, Motohashi and Goto (2010)

講演

特許情報を使った科学技術イノベーション政策研究の事例報告

一橋大学 長岡 貞男

7. おわりに

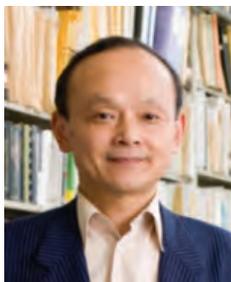
特許情報はイノベーション過程の把握に非常に重要な情報を提供する。

特許情報を補完する基本的なデータを整備することによって、より有効に活用ができる。

また、企業の名寄せや出願人と発明者所属企業の両方の名寄せは、パネルデータ・分析を可能として、企業レベルの要因と発明者レベルの要因を識別することも可能として、研究可能な課題を拡大する。

特許の後方引用はサイエンスからイノベーションへの知識フローを把握する上で重要な役割を果たすと考えられるが、引用が知識フローをどの程度把握するかについて、サーベイデータや事例研究を積み重ねて検証を深める必要がある。引用文献の文献データベースとのリンクは、こうした非常に重要である。事例等による検証・確認が必要(例 Jaffe, Fogarty and Banks (1998)によるNASA 等からのスピルオーバーの研究)である。

発明者の名寄せは、よりチャレンジングであるが、これも発明者のライフサイクルにおける発明活動など動的な分析を可能とする。



長岡 貞男

一橋大学イノベーション研究センター 教授

(経歴)

1975年4月 通商産業省入省

1986年4月 世界銀行へ出向

1992年9月 成蹊大学経済学部 教授

1996年4月 一橋大学商学部附属産業経営研究所 教授

1997年4月 一橋大学イノベーション研究センター 教授

2004年4月～2008年3月 一橋大学イノベーション研究センター長

(現在に至る)

参考文献:

Jaffe, Adam B. Michael S. Fogarty and Bruce A. Banks, 1998, “ Evidence from Patents and Patent Citations on the Impact of NASA and other Federal Labs on Commercial Innovation” , *The Journal of Industrial Economics*, Vol. 46, No. 2, Jun., 1998), pp. 183-205

Jaffe, Adam B., Manuel Trajtenberg, and Michael S. Fogarty. 2002. “The Meaning of Patent Citations: Report on the NBER/Case-Western Reserve Survey of Patentees.” In *Patents, Citations and Innovations*, ed. Adam B. Jaffe and Manuel Trajtenberg, 577-98. Cambridge, MA: MIT Press.

Jones, F. Benjamin (2009) “The Burden of Knowledge and the “Death of the Renaissance Man” :Is Innovation Getting Harder?” , *Review of Economic Studies*, 76, 283-317

Nagaoka, Sadao, Motohashi Kazuyuki and Goto Akira, 2010, “Patent Statistics as an Innovation Indicator,” in Hall, Bronwyn H. and Nathan Rosenberg, eds., *Handbook of the Economics of Innovation*, Vol. 2, Amsterdam : North Holland, 1083-1127

Narin, F. A. Breitzman “1995, Inventive productivity” , *Research Policy* ,24 , 507-519

Thompson Peter “Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations” , *Review of Economics and Statistics*. May 2006, Vol. 88, No. 2, Pages 383-388

Tsukada Naotoshi and Sadao Nagaoka, 2012, “Whether and how does international research collaboration affect invention performance? – Some evidence from triadic patent data of major OECD countries,” Forthcoming IIR Working paper

Walsh P. John and Sadao Nagaoka , 2009a, “How “open” is innovation in the US and Japan?:

evidence from the RIETI-Georgia Tech inventor survey”, RIETI Discussion Paper 09-E-022

Walsh P. John and Sadao Nagaoka , 2009b, “Who Invents?: Findings from the US-Japan Inventor Survey” , RIETI Discussion Paper 09-E-034

長岡貞男、2011、「日米のイノベーション過程：日米発明者サーベイからの知見」藤田昌久・長岡貞男（編）『生産性とイノベーションシステム』第4章、日本評論社

〔研究レビュー 4-2〕

NISTEP における特許情報関連調査研究 のこれまでとこれから

第3調査研究グループ 細野 光章

研究レビュー 4-2

NISTEP における特許情報関連調査研究のこれまでとこれから

第3 調査研究グループ 細野 光章

はじめに

科学技術政策研究所(以下「NISTEP」という。)では、科学技術政策に資するための調査研究を実施してきている。しかしながら、近年、科学技術政策の範囲が拡大されて、社会的・経済的な価値の創出等のイノベーションを促進させるための科学技術イノベーション政策に変質してきている中で、NISTEP における調査研究のあり方も変化・変質しつつある。

NISTEP では、特に基礎的な研究開発活動のアウトプット指標である学術論文に関しては、「サイエンスマップ調査」、「科学技術のベンチマーキング」等で詳細な分析を進めてきているものの、よりイノベーションに近接していると考えられる応用的な研究開発活動のアウトプット指標である特許については、これまでのところ十分な分析をしてきたとは言い難い。

本研究レビューでは、NISTEP における特許及び特許システムに関連する調査研究について、これまでにおいて実施された調査研究を概観するとともに、現在実施中の調査研究の一部を紹介し、更に、将来の調査研究のあり方について考察する。

研究レビュー 4-2

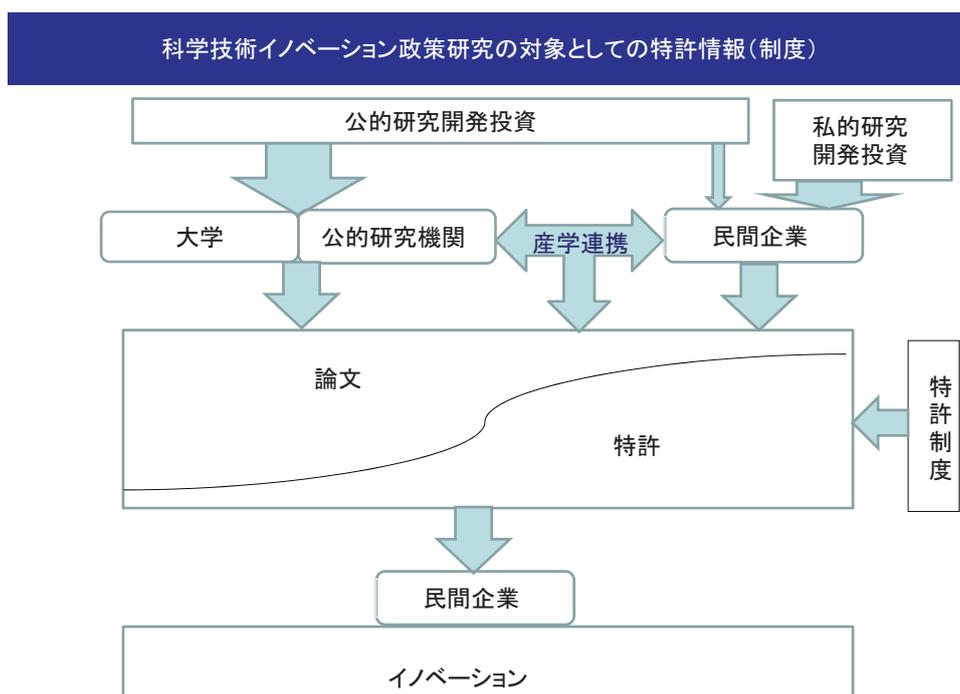
NISTEPにおける特許情報関連調査研究のこれまでとこれから

第3調査研究グループ 細野 光章

1. 科学技術・イノベーション政策研究の対象としての特許情報(制度)

科学技術・イノベーション政策研究の対象としての特許については、大きく二つの見方がある【資料1】。一つは、研究開発の活動指標として特許情報をとらえて行う調査研究であり、公的研究投資及び私的研究開発投資のアウトプット指標としての活用が考えられるほか、研究開発とイノベーションの接合指標としての活用も考えられる。もう一つは、知的財産の適正な管理と活用のバランスを検討するなど特許制度そのものに関する調査研究である。

【資料1】



2. NISTEP におけるこれまでの特許情報(制度)の調査研究

これまでの NISTEP における特許情報・制度に関連する調査研究は、①研究開発のアウトプット指標としての特許情報に関するもの、②民間企業における知的財産活動に関するもの、そして、③大学及び公的研究機関の関連特許に関するもの、と大きく三つに分けることができる【資料 2】。

これら 3 つの方向性がある調査研究に関する NISTEP の代表的な報告書を【資料 2】に示すが、これらのうち特に代表的な研究成果を以下で紹介する。

【資料 2】

NISTEPにおけるこれまでの特許情報(制度)の調査研究

- ・ 研究開発のアウトプット指標としての特許
 調査資料 126「科学技術指標 - 第5版に基づく2006年改訂版」2006年6月
 調査資料 140「科学技術指標 - 第5版に基づく2007年改訂版」2007年7月
 調査資料 144「特許請求項数の国・技術分野・時期特性別分析」2008年1月
 調査資料 155「科学技術指標 - 第5版に基づく2008年改訂版」2008年7月
 調査資料 170「科学技術指標2009」2009年8月
 調査資料 187「科学技術指標2010」2010年7月
 調査資料 198「科学技術指標2011」2011年8月
- ・ 民間企業における知的財産活動
 NISTEP REPORT 135「平成20年度 民間企業の研究活動に関する調査報告」2009年10月
 NISTEP REPORT 143「平成21年度 民間企業の研究活動に関する調査報告」2010年8月
- ・ 大学及び公的研究機関の関連特許(公的研究開発のアウトプット指標)
 調査資料 147「特許出願から見た東北大学の知的貢献分析」2007年9月
 調査資料 154「国立大学法人の特許出願に対する知財関連施策および法人化の影響
 -3大学(筑波大学・広島大学・東北大学)の総合分析-」2008年6月
 調査資料 160「大学および公的研究機関からの特許出願の重点8分野別ポートフォリオ」2008年11月

(1) 研究開発のアウトプット指標としての特許

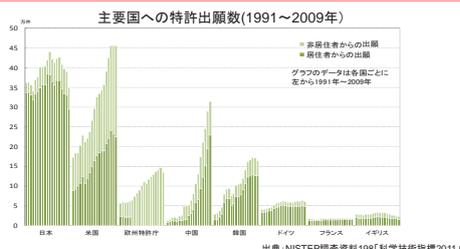
特許制度は国ごとに異なっており、各国・地域の特許庁への特許出願数を単純に比較することはできないが、当該国・地域における特許の重要性及び研究開発活動の活発度を示す指標として捉えることができる。

1991年から2009年までの各国・地域の特許庁への出願件数を見てみると【資料 3】、日本国特許庁(JPO)への出願件数は米国に次ぐ規模であるが、35 万件程度で推移しており、近年は減少傾向にある。米国特許商標庁(USPTO)への出願件数は、1991年以降に倍増して約 45 万件となっているが、近年、横ばいになっている。中国国家知識産権局(SIPO)への出願件数は、年平均成長率約 20%で急増しており、約 31 万件まで増加している。

【資料 3】

研究開発のアウトプット指標としての特許 1

日本への出願件数(約35万件)は、米国に次ぐ規模。
 米国への出願件数(約45万件)は、1991年以降に倍増したが、近年、横ばい。
 中国への出願件数は、年平均成長率約20%で、約31万件まで増加。



注: 1)出願数の内容は、日本への出願を例に取る。以下に対応している。「居住者からの直接出願」日本に居住する出願人が日本特許庁に直接出願したもの、「非居住者からの直接出願」日本以外に居住(例えば米国)する出願人が日本特許庁に直接出願したもの。
 2)国内特許庁の「居住者からの出願」は1996年から値が掲載されていない。
 3)国内移行したPCT出願件数を含む。

研究レビュー 4-2

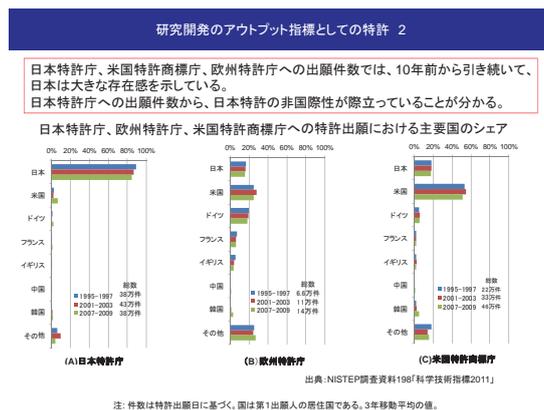
NISTEPにおける特許情報関連調査研究のこれまでとこれから

第3調査研究グループ 細野 光章

日本国特許庁(JPO)、欧州特許庁(EPO)、米国特許商標庁(USPTO)の特許出願における第1出願人の居住国を見てみると(【資料4】)、3極ともに日本に居住する出願人の比率は高く、大きな存在感を示している。

他方で、日本国特許庁への出願人は80%以上が日本国居住であり、我が国における特許の非国際性が際立っているといえる。

【資料4】

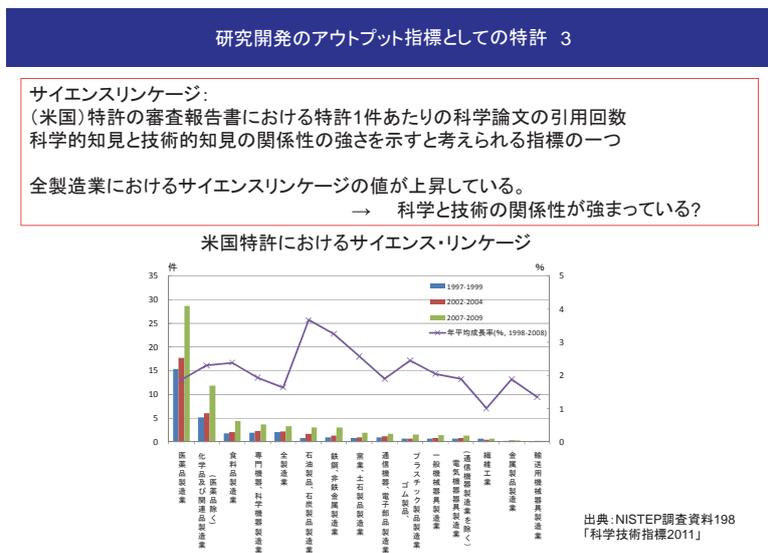


米国特許の審査報告書における特許1件あたりの学术论文の引用回数をサイエンスリンケージ (Science Linkage)と呼び、科学的知見と技術的知見の関係性の強さを示す指標として取り扱われている。

産業分類ごとにこのサイエンスリンケージを見てみると、医薬品製造業のサイエンスリンケージが非常に強く、本産業分類では科学的知見と技術的知見の関係が強固であると考えられる【資料5】。

またサイエンスリンケージの経年変化を見てみると、全製造業において値が上昇しており、科学と技術の関係性が強まっていることが推測される【資料5】。

【資料5】



(2) 民間企業における知的財産活動

特許の出願・維持・活用という知的財産活動は、特にイノベーション創出の主体である民間企業において重要であると考えられるが、それら企業の業種や規模によって知的財産活動のあり方も異なることが推測される。2008 年以降、NISTEP が実施主体となっている「民間企業の研究活動に関する調査」では、民間企業における知的財産活動の状況について調査を行っている。

2009 年度と同調査における業種・資本金階級別の国内特許の出願件数、研究開発費、及び、特許出願性向をしてみると【資料 6】、医薬品製造業は特許出願件数も特許出願性向も低く、非常に特徴的である。これは医薬品製造業における一つの特許の重要性が、他産業と比べて高いことを示していると考えられる。

また、資本金 10 億円以上の 10 億円未満の企業を比較すると、後者の特許出願性向が低く、企業規模が小さいと知的財産活動を行うことが難しくなるということを推測させる。

【資料 6】

民間企業における知的財産活動 1

医薬品製造業は、平均特許出願件数も特許出願性向も低い。1つの特許の重要性が高い。資本金10億円以下と10億円以上の企業で、特許出願性向に差がある。

業種・資本金階級別国内特許保有・自社実施件数(1社平均)

業種	企業数(N)	国内特許出願件数(A)	平均特許出願件数(A/N)	研究開発費(B)(百万円)	特許出願性向(A/B)
建設業	68	1168	17.2	36770	0.032
農林畜産漁業	52	328	6.3	58128	0.099
医薬品製造業	43	329	7.6	123350	0.009
総合化学工業	60	5422	90.4	213290	0.023
その他の化学工業	48	917	19.1	75128	0.012
プラスチック製品製造業	43	2233	52.0	48445	0.049
医薬・土石製品製造業	32	1338	41.8	31152	0.043
鉄鋼業	32	4102	128.2	132287	0.030
精密機器製造業	38	880	23.2	16824	0.041
生産用機械器具製造業	59	3941	66.8	148354	0.027
業務用機械器具製造業	34	3937	115.8	158325	0.037
電子部品・デバイス・電子回路製造業	49	2338	47.7	211980	0.011
その他の電気機械器具製造業	57	10221	179.3	254847	0.040
情報通信機械器具製造業	38	12149	319.7	322371	0.039
自動車・部品製造業	48	3688	76.8	358588	0.010
その他の製造業	39	2657	68.1	72348	0.037
情報サービス業	28	388	13.9	10757	0.037
計	1087	82358	75.8	3177899	0.028

出典：NISTEP REPORT 143「平成21年度 民間企業の研究活動に関する調査報告」

2009 年度と同調査における業種・資本金階級別の国内特許の保有件数及び自社実施件数を見てみると【資料 7】、業種によって特許保有件数に大きな差がみられるが、自社実施率は 50～60%前後と差がないことが分かる。

他方で、資本金規模ごとに比較すると、資本金規模が大きい企業よりも資本金規模が小さい企業の自社実施比率が高くなっている。これは、資本金の小さい企業が厳選して自社の商品化のために特許を活用しているのに対し、資本金の大きな企業は防衛特許等を含む多義的な知的財産活動を行っていることを推測させる。

【資料 7】

民間企業における知的財産活動 2

業種により特許保有件数に大きな差異はあるが、自社実施率に大きな差異は見られない。規模が大きくなるほど特許保有件数が大きく、自社実施率が小さくなる。

業種・資本金階級別国内特許保有・自社実施件数(1社平均)

業種	企業数(N)	特許保有件数(1社平均)	自社実施件数(1社平均)	自社実施率(%) (1社平均)
建設業	67	143.2	45.2	47.3
農林畜産漁業	51	85.5	37.4	32.2
医薬品製造業	43	30.8	12.2	46.0
総合化学工業	54	321.0	209.1	52.8
その他の化学工業	44	107.0	44.2	32.2
プラスチック製品製造業	39	199.0	83.1	52.4
医薬・土石製品製造業	35	179.0	82.5	59.0
鉄鋼業	32	413.3	204.2	59.5
精密機器製造業	35	171.4	97.1	65.2
生産用機械器具製造業	61	327.5	223.5	65.2
業務用機械器具製造業	31	124.8	499.0	57.7
電子部品・デバイス・電子回路製造業	49	199.0	132.8	50.9
その他の電気機械器具製造業	57	816.2	289.4	57.7
情報通信機械器具製造業	38	484.3	267.8	56.9
自動車・部品製造業	44	493.8	237.8	52.4
その他の製造業	39	308.3	113.2	61.7
情報サービス業	20	37.7	17.0	59.7
計	1087	359.6	147.6	55.2

出典：NISTEP REPORT 143「平成21年度 民間企業の研究活動に関する調査報告」

研究レビュー 4-2

NISTEPにおける特許情報関連調査研究のこれまでとこれから

第3調査研究グループ 細野 光章

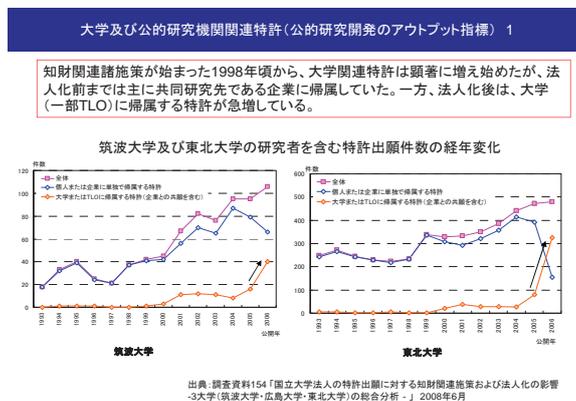
(3) 大学及び公的研究機関関連特許

1990年代半ばから産学官連携によるイノベーション創出が語られるようになったが、特に TLO 法や日本版バйдール法の施行以降、大学及び公的研究機関の関連特許に関する関心が高まり、それら特許を対象とした調査分析が行われるようになってきている。

特に国立大学及び国立研究機関の法人化は、特許管理・活用のあり方に大きく影響を与えたと考えられることから、NISTEP では東北大学、筑波大学、広島大学の協力のもと、当該大学関連特許の法人化前後の特許のあり方について分析を行っている。

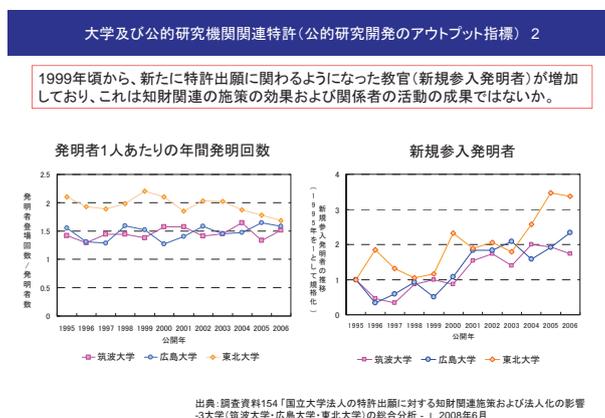
筑波大学及び東北大学の研究者が発明者となっている特許出願数の経年変化を見てみると【資料 8】、大学関連知財施策が始まった1998年頃から、大学関連特許の出願が増加し始め、国立大学の法人化のあった2004年から大学または関連 TLO の特許出願が増加していることが分かる。

【資料 8】



筑波大学、東北大学、広島大学の出願特許の公報情報をもとにした、発明者一人あたりの年間発明回数と新規発明者の経年変化を見ると【資料 9】、1997年頃から新たに特許出願(1999年頃の公開)を行うようになった教官が増加しており、同時期に行われた大学の特許関連施策の影響であると考えられる。

【資料 9】



筑波大学、東北大学、広島大学の出願特許の公報情報をもとにした、出願形態の経年変化を見ると【資料10】、大学が出願人に含まれる特許出願の比率が上昇しており、契約等に基づかないインフォーマルな産学連携活動から権利関係が明確化された契約に基づくフォーマルな産学連携活動に移行しつつあることが推測される。

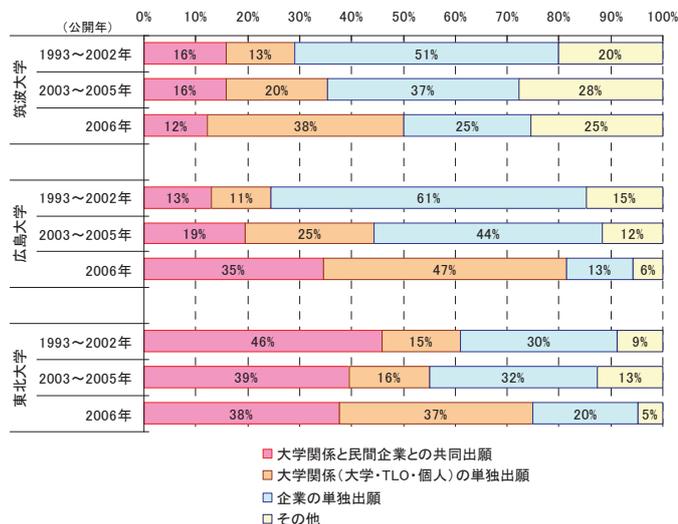
特に国立大学法人化以降は、機関帰属原則となっている職務発明規定等が整備されたことにより、大学が出願人に含まれる特許出願が急増している。

【資料10】

大学及び公的研究機関関連特許（公的研究開発のアウトプット指標） 3

契約等に基づかないインフォーマルな産学連携活動から権利関係が明確化された契約に基づくフォーマルな産学連携活動に移行しているのではないかと推察され、国立大学法人化以降の職務発明に基づく特許の機関帰属化の影響は大きい。

3大学の出願形態の変化



出典：調査資料154「国立大学法人の特許出願に対する知財関連施策および法人化の影響」-3大学（筑波大学・広島大学・東北大学）の総合分析
 年6月

研究レビュー 4-2

NISTEP における特許情報関連調査研究のこれまでとこれから

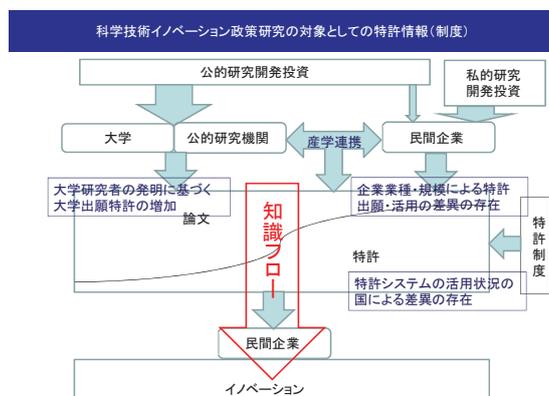
第3 調査研究グループ 細野 光章

3. NISTEP における特許情報(制度)の調査研究の現状とこれから

NISTEP におけるこれまでの特許情報及び制度に関する調査研究では、前述したように①研究開発のアウトプット指標としての特許情報に関するもの、②民間企業における知的財産活動に関するもの、そして、③大学及び公的研究機関の関連特許に関するものがあり、関連の知見を得ているが【資料 11】、それらは端緒についたばかりであり、更なる調査研究が必要である。

特許情報及び特許制度を対象とした調査研究には、その対象やアプローチの違いで様々なものがあると考えられるが、現在、NISTEP では特に公的研究開発投資からイノベーション創出への知識フローのあり方を調査研究しており、そのような知識フローを明示化・可視化するための情報として特許を活用できないか、検討を進めている。

【資料 11】



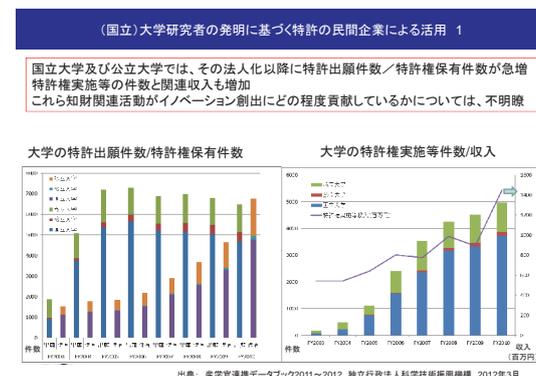
(1) NISTEP における特許情報(制度)の調査研究の現状

1990 年代半ば以降、イノベーション創出を目的として産学官連携とそれに付随する知的財産活動を推進するための施策が実施され、それに呼応するように大学の特許出願件数、特許保有件数、特許権実施件数、及び、特許権実施料収入も増加してきている【資料 12】。

しかしながら、このような大学における知的財産活動が企業等におけるイノベーション創出にどの程度寄与しているか、即ち、大学から企業等への知識フローがどのようになっているかについては、一部の事例研究を除き、我が国においては体系的な調査研究が実施されておらず、十分な知見が得られていない。

このため、現在、NISTEP では特に大学から企業への知識フローの指標として大学関連特許に着目し、その調査研究を進めている。

【資料 12】



大学関連特許の分析を行うにあたり、当該特許のデータベースの構築が必要である。このため、NISTEP では、企業等との共同研究等の実績があり、また、職務発明の機関帰属がほぼ原則化している国立大学法人を対象としてデータベース(以下、「国立大学関連特許データベース」という。)を構築した【資料 13】。

なお、国立大学関連特許データベースの構築にあたっては、国立大学法人の研究者(教職員、学生)が発明者である特許を特許公報データベースから抽出し、出願人・発明者の名寄せ等を行うとともに、企業規模・業種や大学規模・類型等の関連属性を付加している。

【資料 13】

(国立)大学研究者の発明に基づく特許の民間企業による活用 2

2004年～2007年度に出願された、国立大学、または国立大学に所属する職員、学生等が発明者として関与する特許は、20,485件
発明者の名寄せ・同定作業が非常に困難(同姓同名、所属異動、原本記載ミス等)

国立大学関連特許データベースの構築

大学関連特許の抽出にあたっては、発明の機関帰属がある程度徹底しており、いわゆる研究大学として取り扱うことが可能な国立大学(83校)を対象とすることとした。また、機関帰属が開始されたと考えられる国立大学の法人化以降の2004～2007年度に出願された特許を対象とした。

(1)抽出方法

- ①「出願人」または「発明者住所」に国立大学名称が記載される特許
- ②「出願人」にTLOが含まれる特許のうち、発明者に国立大学研究者が記載された特許
- ③「出願人」に(独)科学技術振興機構(以下、JSTと略す)が含まれる特許のうち、発明者に国立大学研究者が存在する特許

(2)情報の補完

- 前記②および③の抽出において、発明者の住所は発明者個人の居所が記載され、所属が不明である特許が殆どである。このため、発明者の出願時の所属は、全国大学職員録や、KAKEN研究者検索などのインターネット情報を利用して特定し、情報を補完している。また、同姓同名者、および同一人の異なる所属での発明はこの過程で区分している。ちなみに、データベースに含まれる発明者の総数は33000人余りである。
- さらに、分析には国立大学と共同で研究を実施する企業の属性情報を必要とする。データベースに含まれる出願人または発明者が所属する約3000社の企業の資本金、従業員数、業種(主業)等の属性情報は、帝国データバンク(株)より購入、または企業ホームページ等から入手することにより情報の付加を行っている。

研究レビュー 4-2

NISTEPにおける特許情報関連調査研究のこれまでとこれから

第3調査研究グループ 細野 光章

大学から企業等への知識フローについて、大学関連特許を活用して分析するにあたり、当該知識フローのあり方を明示的に規定する我が国の特許法を見ておく必要がある。

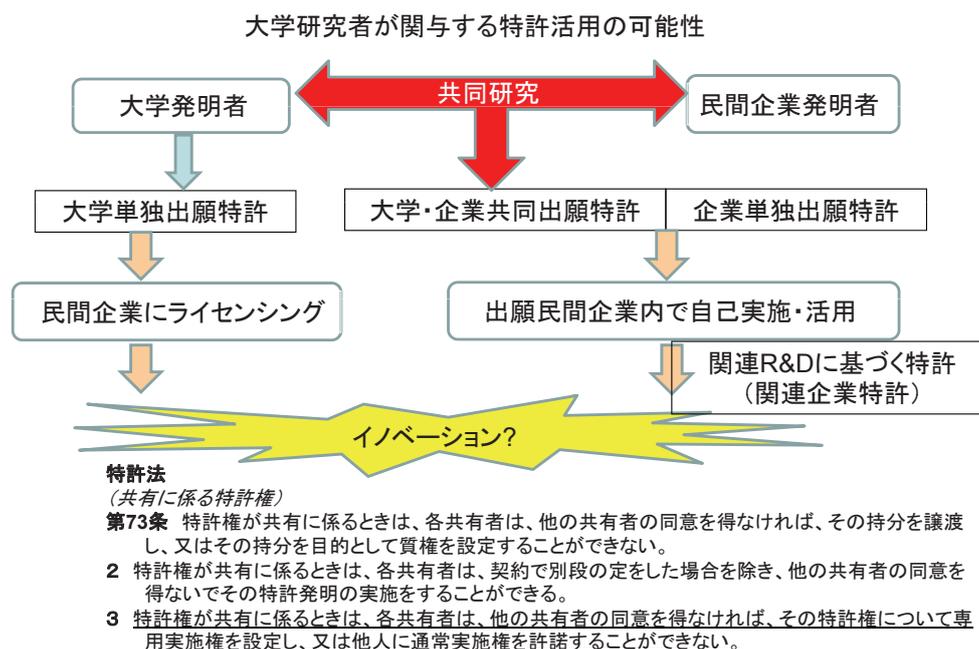
大学発明者が関与する特許は、大学発明者による発明を当該発明者の所属大学が単独出願した特許(大学単独出願特許)、そして、大学研究者及び企業等研究者による共同発明を当該発明者の所属大学及び企業等が共同で出願する出願特許(大学・企業共同出願特許)に大別される【資料14】。

我が国の特許法では共有に係る特許権について、「特許権が共有に係るときは、各共有者は、他の共有者の同意を得なければ、その特許権について専用実施権を設定し、又は他人に通常実施権を許諾することができない。」と規定しており、別に定めのない限り大学・企業共同出願特許は、共同出願をしている企業の同意なく大学は第三者に実施許諾できない。

したがって、大学・企業共同出願特許は、共同出願をしている企業が自己実施しない限り、当該特許の積極的な活用が見込めない、すなわち、大学で創出された知識フローが共同発明先企業内で止まったといえる。

【資料14】

(国立)大学研究者の発明に基づく特許の民間企業による活用 3



先に構築した国立大学関連特許データベースをもとに、その出願人及び発明者の属性を分析したところ【資料15】、国立大学及び関連TLOの単独出願特許は全体の40.7%であり、国立大学・民間企業の共同出願特許が30.7%、民間企業の単独出願特許が7.4%であった。

前述したように大学・企業共同出願特許の第三者実施許諾の困難さを考慮すると、企業単独出願及び大学・企業共同出願特許(38.1%)は、当該出願特許企業による自己実施等がなければ、活用の余地がほとんどない。

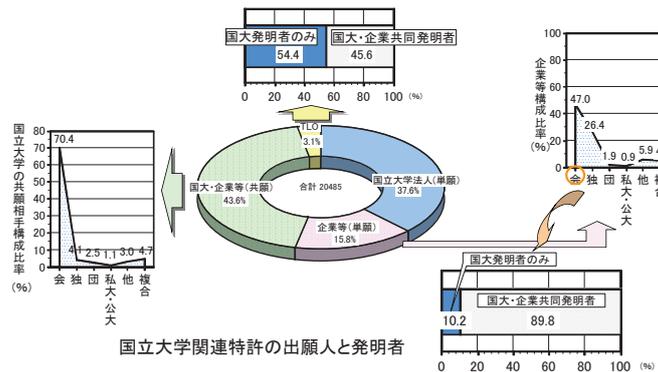
研究レビュー 4-2
NISTEP における特許情報関連調査研究のこれまでとこれから
第3 調査研究グループ 細野 光章

このため、大学から企業への知識フローのあり方を調査するために、その大きな流れとなっている大学・企業共同出願特許の出願企業内における当該共同出願特許の活用について、より詳細な調査研究を行う必要がある。

【資料 15】

(国立)大学研究者の発明に基づく特許の民間企業による活用 4

国立大学及びTLOによる単独出願は全体の40.7%
国立大学研究者が発明者に含まれる企業単独出願特許は全体の7.4%
企業単独出願及び企業・国大との共同出願特許(38.0%)は、当該出願特許企業による自己実施等がなければ、活用の余地がほとんどない。



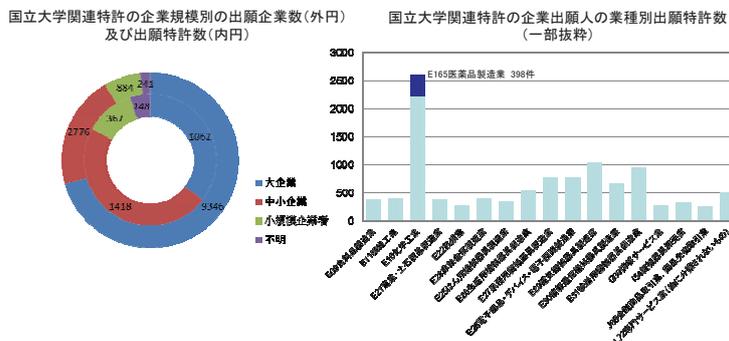
国立大学関連特許データベースをもとに、出願人に民間企業が含まれる特許について、当該企業の属性(企業規模及び業種)を分析したところ【資料 16】、出願特許数の70.6%に大企業が関与しており、産学連携活動における大企業のプレゼンスの大きさが明らかになった。

国立大学関連特許の企業出願人の業種ごとに特許出願件数をみると、化学工業が突出している。他方、日本国特許庁が実施している出願件数上位 301 社を対象にした調査では、電気機械器具製造業の特許出願数が突出している。産学連携活動においては化学工業に属する企業のプレゼンスの高さが明らかになった。

【資料 16】

(国立)大学研究者の発明に基づく特許の民間企業による活用 5

国立大学関連特許の企業が出願人に含まれる特許のうち、大企業が70.6%に関与。また、出願業種別特許出願数では、化学工業が突出している。特許庁調査(出願上位301社を対象)では、電気機械器具製造業の出願特許数が突出している。



研究レビュー 4-2

NISTEPにおける特許情報関連調査研究のこれまでとこれから

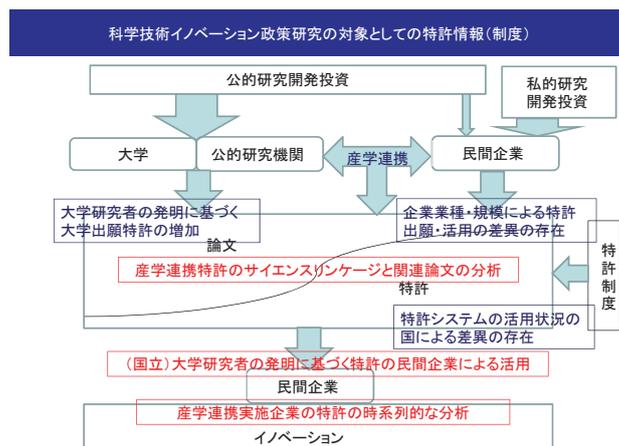
第3調査研究グループ 細野 光章

(2) NISTEPにおける特許情報(制度)の調査研究のこれから

前述したように、現在、NISTEPでは大学研究者の発明の民間企業での活用のあり方について調査研究を実施中である。具体的には、大学・企業共同出願特許及び関連する大学単独出願特許及び企業単独出願特許を対象として、その関係性を書誌情報学的に分析すると共に、大学・企業共同出願特許の大学及び企業発明者に対する大規模アンケート調査を実施している。

これに加え、将来的に NISTEP では、産学連携特許のサイエンスリンクージと関連論文の分析と産学連携実施企業の特許の時系列的な分析を実施する必要がある【資料 17】。

【資料 17】

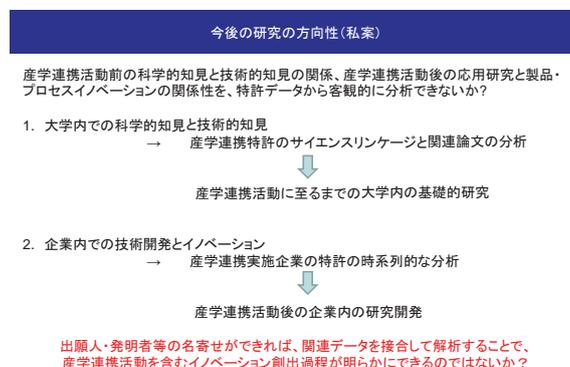


産学連携活動に至るまでの大学内の基礎的研究のあり方については、産学連携特許のサイエンスリンクージと関連論文の分析することにより、また、産学連携活動後の企業内の研究開発のあり方については、産学連携実施企業の特許を時系列的に分析することにより、ある程度、把握することが可能であると考えられる【資料 18】。

これらの調査研究と現在実施中の調査研究の結果を統合し、産学連携活動前の科学的知見と技術的知見の関係、産学連携活動後の応用研究と製品・プロセスイノベーションの関係性を、特許データから客観的に分析することが可能になるかもしれない。

ただし、このような調査研究を実施するためには、特許情報のうち出願人・発明者等の名寄せ等のデータクリーニングが必須であり、この点が大きな課題といえる。

【資料 18】





細野 光章

第3 調査研究グループ 上席研究官

(経歴)

1993年3月 名古屋大学大学院農学研究科修士課程修了
1998年8月 アジア経済研究所開発スクール修了
1999年11月 インペリアルカレッジ環境技術センター修士課程修了
2003年3月 筑波大学大学院ビジネス科学研究科修士課程修了
この間、民間企業、NGO、国立研究所等に勤務
2004年8月 東京工業大学産学連携推進本部 特任准教授
2010年4月 文部科学省科学技術政策研究所 上席研究官
(現在に至る)

参考文献:

- 調査資料-126 「科学技術指標 - 第5版に基づく2006年改訂版」(2006)
調査資料-140 「科学技術指標 - 第5版に基づく2007年改訂版」(2007)
調査資料-144 「特許請求項数の国・技術分野・時期特性別分析」(2008)
調査資料-147 「特許出願から見た東北大学の知的貢献分析」(2007)
調査資料-154 「国立大学法人の特許出願に対する知財関連施策および法人化の影響
-3 大学(筑波大学・広島大学・東北大学)の総合分析-」(2008)
調査資料-160 「大学および公的研究機関からの特許出願の重点8分野別ポートフォリオ」
(2008)
調査資料-155 「科学技術指標 - 第5版に基づく2008年改訂版」(2008)
調査資料-170 「科学技術指標 2009」(2009)
調査資料-187 「科学技術指標 2010」(2010)
調査資料-198 「科学技術指標 2011」(2011)
NISTEP REPORT No.135 「平成20年度民間企業の研究活動に関する調査報告」(2009)
NISTEP REPORT No.143 「平成21年度民間企業の研究活動に関する調査報告」(2010)

文部科学省 科学技術政策研究所

〒100-0013 東京都千代田区霞が関 3-2-2

TEL:03-3581-2466 FAX:03-3503-3996

<http://www.nistep.go.jp>

2013年3月



<http://www.nistep.go.jp>