

Title	サイエンス・リンケージ手法を用いた科学依拠型産業の時系列分析
Author(s)	玉田, 俊平太; 内藤, 祐介; 玄場, 公則
Citation	年次学術大会講演要旨集, 29: 653-658
Issue Date	2014-10-18
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/12533
Rights	本著作物は研究・技術計画学会の許可のもとに掲載するものです。This material is posted here with permission of the Japan Society for Science Policy and Research Management.
Description	一般講演要旨

サイエンス・リンケージ手法を用いた科学依拠型産業の時系列分析

○玉田俊平太（関西学院大学）、内藤祐介（人工生命研究所）、玄場公則（立命館大学）

科学研究の成果がイノベーションに与えている影響を定量的に計測する手法の1つとして、特許本文中に引用されている科学論文を調べる「サイエンス・リンケージ」という手法がある。本研究では、構築するとともに、機械学習アルゴリズムを用いた1993年以降に特許公報に掲載された日本特許の全数・全文解析により、特許本文中に引用されている論文を自動抽出し、それにより計測された特許の科学依拠度が、①分野ごとにどのような違いがあるのか、②イノベーションが科学に依拠する度合いは年とともに増えているのか減っているのか、などといった問いに対して調査を行った。

1. 用いた手法

発表者らは以前、科学技術基本計画の重点4分野に属する特許に引用されている論文等（科学依拠度：サイエンス・リンケージ）を研究チームの人間が目視により計測した結果を発表した。しかしながら、年間20万件以上が成立する特許の引用論文等をに対してこのアプローチを適用することは、膨大な時間と人手が必要で、大学に所属する研究者には事実上不可能である。

この壁を突破するためには、特許が引用している論文などを抽出する作業を、何らかの形で自動化する必要がある。しかし、論文に他の論文を引用する場合と同様、特許明細書中に参照論文を記載する際の書式は統一されておらず、単純な検索アルゴリズムでは引用されている論文を抽出し損なったり、逆に論文ではない文字列を論文と誤認してしまったりする。

この問題を解決するため、前章で述べた、人間が目視によって抽出した、特許に引用されている論文の情報（文字列）を「教師」として用い、有限状態機械アルゴリズムを用いた独自のソフトウェアの開発を行った。そのソフトウェアを用いて、1995年から2013年までに特許公報に掲載された全特許、約239万6千件を対象とし、各特許一件一件の中に引用されている論文および特許の数を自動で計測することに成功した。（詳細については拙著『産学連携イノベーション—日本特許データによる実証分析』を参照のこと）

2. 結果

特許化されたイノベーションは、国際特許分類(IPC: International Patent Classification)に基づき、各特許につき一つ以上の技術分類が割り当てられている。それらのうち最初に付与されている技術分類を筆頭IPCという。国際特許分類はすべての技術が8つのセクション(A 生活必需品、B 処理操作；運輸、C 化学；冶金、D 繊維；紙、E 固定構造物、F 機械工学；照明；加熱；武器；爆破、G 物理学、H 電気)のいずれかに分類され、それぞれのセクションはさらに複数のクラスに細分され、それぞれのクラスがさらに細かいサブクラスへと分かれている。

つまり、国際特許分類を用いれば、特許化されたイノベーションを、それがどのような技術に属しているかに基づいて分析することができるのだ。

もし、異なる技術分類に属する特許に引用されている論文の数が、互いに大きく異なるのであれば、それは発明者の頭にアイデアがひらめいた際にどの程度科学的知識を用いたかが、技術の種類によって大きく異なっていることを示すことになるだろう。

そこでまず、我々が保有している特許データベースの中の特許全てについて、一件毎に、その特許に引用されている論文(及び先行特許)を自動的に抽出し、その数をカウントした。対象とした特許は1993年から2013年の20年間に特許公報に掲載された239万8千254件である。

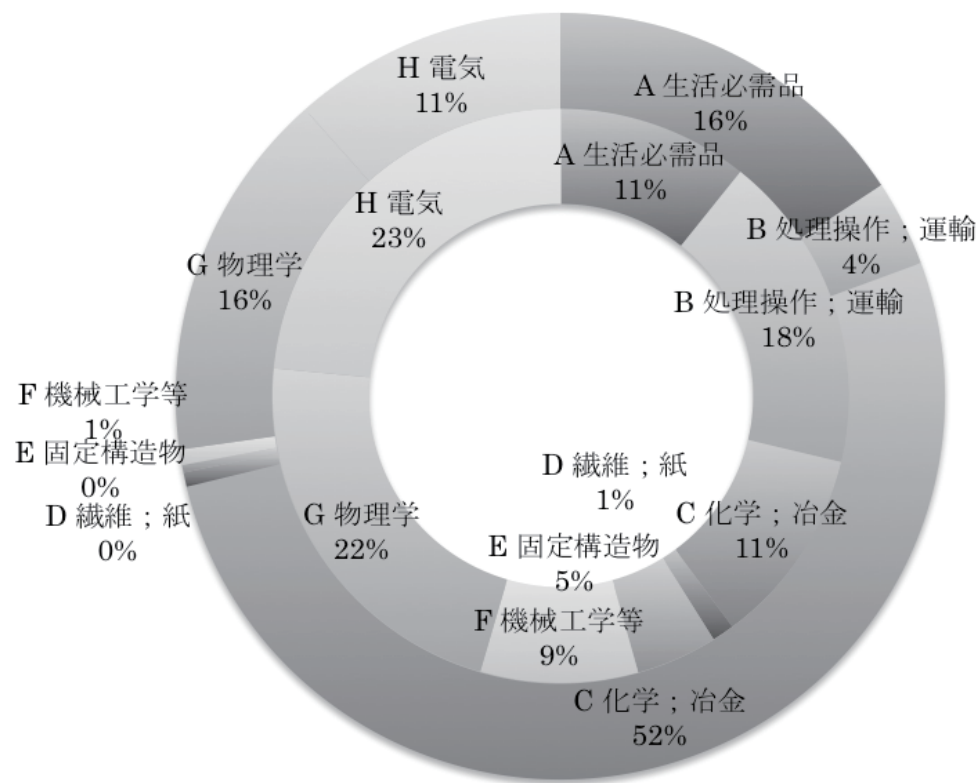


図 1 技術分類別特許比率およびサイエンス・リンケージ比率

図 1 の内側の円グラフは、240 万件の特許が 8 種類の国際特許分類のいずれに分類されているのかを、外側の円グラフは特許に引用されている論文がどの技術分類に属する特許に引用されたものであるかについて、それぞれの百分率を示したグラフである。

たとえば、「A 生活必需品」に分類される特許は全特許の約 11% であるが、そこに分類された特許に引用されている論文の数は、すべての特許に引用された論文の数の合計の約 16% を占める。すなわち、「A 生活必需品」に分類された特許の比率より、A 分野特許に引用された論文が全論文に占める割合の方が高いため、「A 生活必需品」に分類された特許は、相対的に科学の知識に依拠している割合が高いとすることができるだろう。

逆に「B 処理操作；運輸」に分類される特許については、全特許の 18% を占めるにもかかわらず、そこに引用されている論文は全体の引用論文の 4% にすぎない。すなわち「B 処理操作；運輸」分野の特許は、相対的に、科学知識に依拠している度合いが少ないとすることができよう。

そして、最も科論文の知識に依拠している度合いが大きいのが、「C 化学；冶金」の分野である。この分野は全特許のわずか 11% を占めるに過ぎないにもかかわらず、それらの特許に引用されている論文が全引用論文に占める割合は 52% をも占める。すなわち、「C 化学；冶金」に属する特許は、通常の 5 倍も科学の知識に依拠している度合いが高いのだ。そして、特許全体の 22% を占めるに過ぎない「A 生活必需品」と「C 化学；冶金」の 2 つの技術分類に引用されている論文だけで、全特許に引用されている論文の実に約 7 割を占めるのである。

逆に、「D 繊維；紙」、「E 固定構造物」、「F 機械工学；照明；加熱；武器；爆破」の三つの分野の特許は、ほとんど科学論文を引用しておらず、これらの特許は、科学の知識はほとんど頼りにしていないと思われる。

表 1 サイエンス・リンケージの多い分野トップ 20

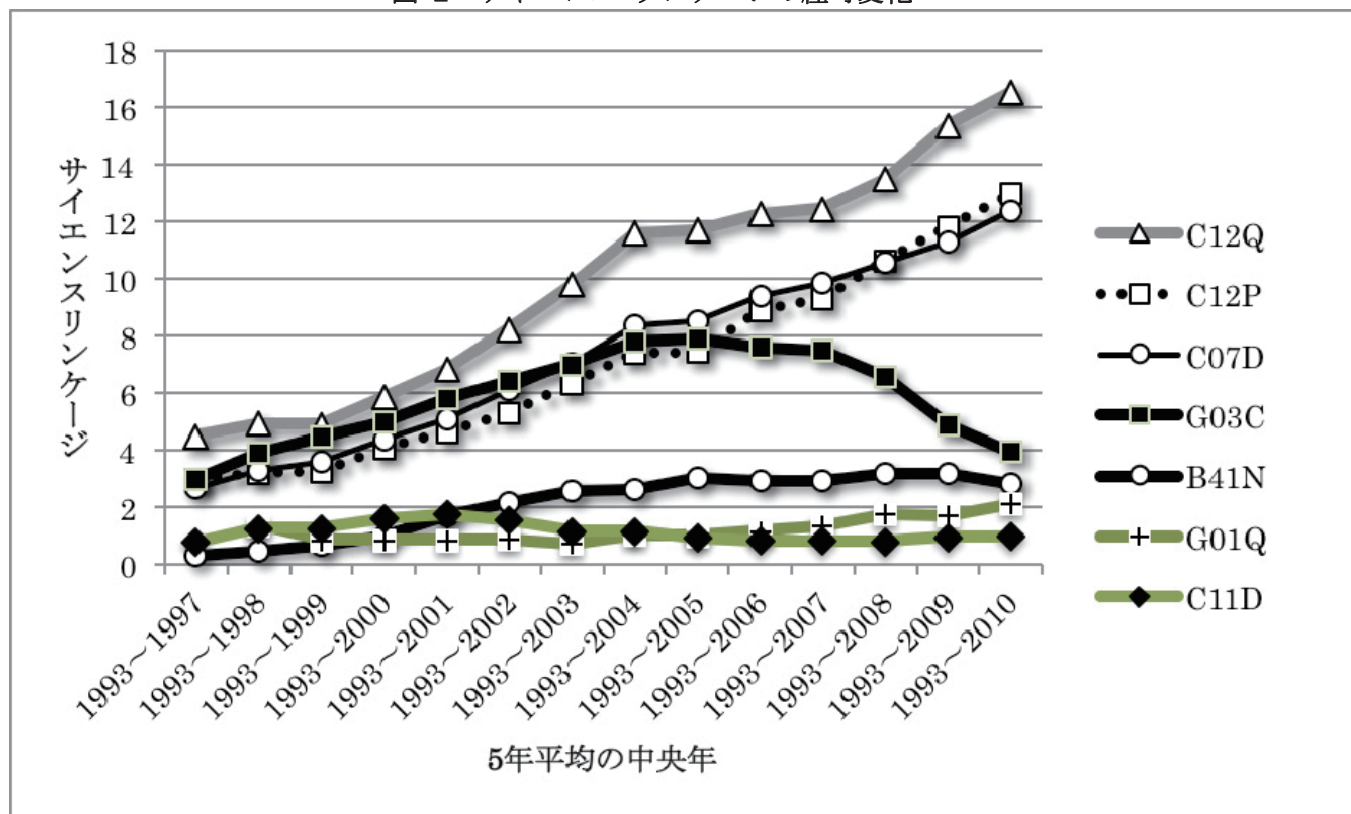
順位	IPC 分類	説明	論文 引用数	特許数	サイエンス・リンケージ
1	C40B	コンビナトリアルケミストリ；ライブラリ，例．ケミカルライブラリ，コンピュータ内でのライブラリ	439	22	20.0
2	C12N	微生物または酵素；その組成物；微生物の増殖，保存，維持；突然変異または遺伝子工学；培地	272476	14003	19.5
3	C07K	ペプチド	80463	4332	18.6
4	A61P	化合物または医薬組成物の治療活性	440	28	15.7
5	A01H	新規植物またはそれらを得るための処理；組織培養技術による植物の増殖	4871	374	13.0
6	C12Q	酵素または微生物を含む測定または試験方法そのための組成物または試験紙；その組成物を調製する方法；微生物学的または酵素学的方法における状態応答制御	35139	3326	10.6
7	C07D	複素環式化合物	134900	17354	7.8
8	C07J	ステロイド	3469	472	7.3
9	C12P	発酵または酵素を使用して所望の化学的物質もしくは組成物を合成する方法またはラセミ混合物から光学異性体を分離する方法	18592	2626	7.1
10	C07H	糖類；その誘導体；ヌクレオシド；ヌクレオチド；核酸	10027	1501	6.7
11	A61K	医薬用，歯科用又は化粧品用製剤	153864	29146	5.3
12	C12S	酵素または微生物を利用して既存の化合物または組成物を遊離，分離または精製する方法；酵素または微生物を利用して繊維製品を処理するまたは材料の固体表面を清浄する方法	42	10	4.2
13	G03C	写真用感光材料	11009	2625	4.2
14	G21G	化学元素の変換；放射線源	313	88	3.6
15	B82B	ナノ構造物；その製造または処理	1187	337	3.5
16	C07F	炭素，水素，ハロゲン，酸素，窒素，硫黄，セレンまたはテルル以外の元素を含有する非環式，炭素環式または複素環式化合物	13297	3888	3.4
17	G06J	ハイブリッド計算装置	13	4	3.3
18	C08B	多糖類，その誘導体	3375	1110	3.0
19	C07C	非環式化合物または炭素環式化合物	47712	17715	2.7
20	G06N	特定の計算モデルに基づくコンピュータ・システム	1433	716	2.0

表1からわかるのは、前節の、セクション毎のサイエンス・リンケージ分析でも示唆されていたとおり、科学の知識を引用している特許はバイオ関連分野のものが大半である、という事実だ。実際、サイエンス・リンケージの多い技術分野トップ20のうち、1位から12位までがバイオ関連分野である。

その他の分野では、13位にやっと写真用感光材料（ナノテク）、14位に化学元素の変換；放射線源、15位にナノ構造物、17位にハイブリッド計算装置、20位に特定の計算モデルに基づくコンピュータ・システムが現れる。これらは大半がナノテクかIT分野だと考えられる。

バイオ分野、ナノテク、IT分野などにおいてサイエンス・リンケージが特に強いということは、そういった分野の発明を発明者が思いついた際に、頭の中の多くの科学的知識に依拠していたということを示していると考えられる。このことは、バイオ分野やナノテク、IT分野などがサイエンス型産業と呼ばれ、世界の有力大学の周辺にバイオベンチャーが生まれてクラスターを形成していることと整合的である。今後、産学連携政策や地域クラスター等の地域振興政策を立案する際には、技術分野によってサイエンス・リンケージが非常に大きく異なるという事実を踏まえ、技術分野ごとの特性を踏まえたよりきめ細かな政策の作り込みが求められよう。

図2 サイエンス・リンケージの経時変化



これ我々は、特許データベースの中にある約20年分の特許を、技術分類という1つの軸だけで切って分析してきた。しかし、科学とイノベーションの結びつきは時間が経っても変わらないのだろうか？イノベーションにライフサイクルがあるように、サイエンス・リンケージにもライフサイクルがありえるのではないだろうか？

こうした問いに答えるため、サイエンス・リンケージの経時変化を調べた。図2は、2006年から2010年までの5年間平均サイエンス・リンケージの値が、その10年前の1996年から2000年までの5年間平均サイエンス・リンケージと比べて、最も伸び率が高かった上位1位から5

位までの IPC 分類（いずれも白色のマーカー）と、逆に、10 年前と比べてサイエンス・リンケージの伸び率がマイナスになった技術分類 2 つ（いずれも黒色のマーカー）を、時系列でグラフ化したものである。

2 の横軸は、特許出願年の 5 年間移動平均期間を示している。縦軸は、当該技術分類の平均論文引用数である。ただし、1993 年から 2013 年までの約 20 年間の特許件数の合計が 300 件に満たない技術分類は年ごとのサイエンス・リンケージ増減の変動が大きくなるため除外してある。また、2011 年以降に出願された特許は、そもそも成立している特許の数が極端に少なくなるため除外してある。

図 2 の中で、サイエンス・リンケージの絶対値が一番大きいのは C12Q だが、サイエンス・リンケージの伸び率が最も大きかったのは絶対値では 2 番目の「C12P 発酵または酵素を使用して所望の化学物質もしくは組成物を合成する方法またはラセミ混合物から光学異性体を分離する方法（伸び率 3.20）」であった。次いでサイエンス・リンケージの伸びが大きかったのは「B41N 印刷版またはフォイル；印刷，インキ付け，湿し等，印刷機に使用される表面用材料；その表面の使用準備または保存（伸び率 2.94）」であり、その次が「C07D 複素環式化合物（伸び率 2.85）」、さらに「C12Q 酵素または微生物を含む測定または試験方法；そのための組成物；そのような組成物の製造方法（伸び率 2.82）」、「G01Q 走査プローブ技術または装置；走査プローブ技術の応用，例．走査プローブ型顕微鏡（伸び率 2.53）」と続く。サイエンス・リンケージの伸び率が大きい分野も、その多くがバイオテクノロジー分野であり、一部ナノテクノロジーが含まれていることが明らかとなった。ちなみに特許全体のサイエンス・リンケージは、10 年間で 2.12 倍に増えていた。

逆に、サイエンス・リンケージが 10 年前より一番減少した分野は「C11D 洗浄性組成物；単一物質の洗浄剤としての使用；石けんまたは石けん製造；樹脂石けん；グリセリンの回収（伸び率 0.60）」であった。減少率で下から 2 番目だった分野は「G03C 写真用感光材料；写真法，例．映画，X 線写真法，多色写真法または立体写真法；写真の補助処理法（伸び率 0.79）」であった。

サイエンス・リンケージの伸びが大きかった技術分野はなぜそうなったのか、逆に、サイエンス・リンケージが減少した技術分野はなぜ減少したのかを解明するにはまだまだ追加的な研究が必要である。しかし、少なくともイノベーションが科学の知識に依拠する度合いが時間の経過とともに変化するというエビデンスは得られた。

社会や個人のニーズの変化に応じて、求められる製品やサービスが変化し、それに伴ってそうした製品やサービスを実現するために求められる科学的知識も変化することはこれまでもしばしば起きてきたし、今後も起きるであろうことは十分に予測できる。一例を挙げれば、30 年程前にはエレクトロニクスを学ぶものの必須科目であったアナログ回路の設計技術は、ありとあらゆるもののデジタル化によって、今や、その必要性は大幅に減少していると言わざるを得ない。

アナログ回路を支えてきた三角関数などの数学や電磁気学や固体物理学などの知識も、デジタル回路の設計には必ずしも必要とされない。昔は真空管やトランジスタ、トランス、コイル、コンデンサーなどのアナログ部品で構成されていたラジオやテレビも、今ではデジタル回路が取って代わり、一つのチップの上に実装され、さらに最近ではソフトウェアによって同じ機能が実現されるに至っている。

このように、社会のニーズの変化に応じて、どのような科学の知識を学ぶことが必要であるかも、自ずと変化してきていると言えるだろう。

3. 考察

以上の調査結果から導き出される政策提言は以下の通りである。イノベーションの分野別サ

イエンス・リンケージ調査が示唆するのは、科学とイノベーションとのリンケージの強さがその分野によって大きく異なっているということである。そして、特許中に引用された論文を指標とした科学と技術のリンケージの強さは、バイオテクノロジー分野が突出していた。

しかし、他の技術分野でも、ナノ構造物、ハイブリッド計算機、流体回路素子、走査プローブ、暗号、光コンピューティング、音声合成・認識などに関連する技術などが科学との強い連関を示した。もとより、科学の目的はイノベーションのみにあるわけではない。したがって、社会からのニーズなど他の要素も勘案しつつ、政府はこうしたイノベーションとの結びつきの強い分野に注力することを通じて、国民のゆとりと豊かさを目指すべきではないだろうか。

また、産学連携を行う場合には、こうした分野による科学との結びつきの強さの違いを十分に念頭に置いてやり方を変える必要があるだろう。例えば医薬品やIT産業分野の一部などの、特許の占有可能性が高くかつ製造の規模が小さくて済み製造コストも比較的低い産業分野においては、特許のライセンスによる技術の移転も可能であろうし、それをてこにした大学発ベンチャーもまた可能であろう。一方、製品の複雑性が高く、規模の経済が働き、サービス網の整備や販売網などイノベーションの補完的資産の維持が不可欠な航空・宇宙や自動車などの産業分野は、やはり大企業の独壇場であり産学連携によるメリットは一部に限られるだろう。

また、プロセス・イノベーションが重要な分野は、特許によって技術を公開しライセンスするより、秘密保持契約の締結と共同研究によって競争力維持・向上が期待できよう。

一部の大学職員や研究マネージャーの中には、こうした技術を競争力に結びつける方法に複数のやり方があることを知らず、いたずらに特許出願数の増大のみを目標として、プロセス技術であってもどんどん特許化してしまっている例があるという。これでは、敵に塩を送っているに等しい。

4. 限界

最後に、サイエンス・リンケージを用いた研究手法の限界について述べる。前述したように、サイエンス・リンケージとは、特許化されたイノベーションの中に、学術論文等の科学的知識がどの程度引用されているかを計測した指標である。これによって解るのは、特許中に引用されている論文についての情報のみである。

例えば、高校や大学・大学院で習う、教科書に掲載されているような基礎的な科学の知識、例えば、微分方程式の解法であるとか、半導体の基礎理論、生化学における遺伝子操作の原理などは、特許の新規性を主張する際に貢献するものではないため、わざわざ発明者が特許中に引用する事は考えにくい。発明者によって特許に引用されるのは、特許の新規性を主張する際にそのサポーティング・エビデンスとなるような最新の学術研究の成果であったり、そのデータを取得する際に用いられた先進的な計測手法などであるということが、筆者があるバイオ企業の研究所長クラスの方々に行ったアンケート調査によって明らかとなっている。

また、大学や公的研究期間ではなく、企業において活発に研究が行われている分野では、きわめて基礎的・基盤的なアイデアでも論文として発表せずに、まず特許として出願されると聞いている。そうした場合、特許を出願する時点では、引用すべき科学論文はそもそも存在せず、結果としてそうした分野のサイエンス・リンケージは低くなってしまう可能性がある。

つまり、本研究で用いた、特許中に引用されている論文数を計測することによって、イノベーションにどの程度科学的知識が貢献しているかを計測する方法は、イノベーションに対する基礎的な科学知識の貢献を過小に評価してしまう可能性があることに留意が必要である。