

## 発明者引用特許の抽出とその分析

六車正道\*

発明者引用特許の抽出とその分析 六車正道【主な著作、講演】  
情報の科学と技術(情報科学技術協会) 57巻、7号(2007)

\*むぐるま まさみち 六車技術士事務所

所在地 〒316-0032 日立市西成沢町4-37-6

電話 050-8012-2416, Fax 050-8007-7906

支店 〒132-0024 東京都江戸川区一之江6-10-14-609

Tel 050-8012-2416(原稿受領2007426)

日本特許における引用文献の種類と特徴を解説し特許明細書中に発明者自身が書いている引用特許の記載の特徴を明らかにした。そして、「発明者引用特許」の高精度の抽出システムを完成させ、1993年以降の発明者引用特許の記載の実態を明らかにし、年々、記載件数が増えている状況を明らかにした。また、一例としてナノテクノロジー分野において審査引用特許に対して発明者引用特許が圧倒的に有利であることを明らかにした。

### 1. はじめに

特許を含む文献の引用情報は知識の断片である個々の文献を知識の流れとして体系付ける優れた資源であり、発明を含む知識の再生産に貢献する。

特許においては、審査請求された特許は、特許庁の審査による引用情報が公開される。その外に、発明者が特許出願の明細書中に自ら記載した引用文献(以後、発明者引用という)がある。発明者引用は出願から1年半後の公開公報によりすべてオープンになり、審査引用にくらべて全件が公開され、時期的に早く、さらに技術的な広さの面でも優れている。さらに、一般文献の引用に対して引用特許は特許番号で統一されており様々な利用が期待される。しかし、審査引用に比べて発明者引用の特許番号のフォーマットは一定しておらず取り扱いには工夫を要する。

本報告では、日本特許における各種の引用特許の特徴を検討した。さらに、発明者引用特許の約99.99%の自動抽出に成功したのでその実態を報告し、ナノテクノロジー分野における発明者引用特許の詳細な状況を報告する。

### 2. 引用特許のこれまでの利用

特許の引用データの分析は、審査官による引用が古くから数多く特許公報に掲載されてきた米国特許において進められてきた。特に、ある特許が登録された後、他の特許に引用される「被引用」データは特許の客観的な評価指標として注目を集めてきた。1970年代から米国のCH社による審査引用特許の分析は先駆的であり、わが国でもゼファー社から翻訳論文集1)が発行されている。

また、最近ではオンラインとパソコンを利用した分析システムがトムソン社(Aureka)、ウィズドメイン社(Focust)、WIPS社(WIPS)などが市販されている。これらは単純な引用、被引用だけでなく二次引用や共引用などの間接引用も利用できる。また、FocustやWIPSでは引用・被引用で集めた集合を対象にキーワードや特許分類、出願人で絞り込むことができ、元の特許とさらに類似性の高い特許に絞り込むことが可能2)としている。これらはいずれも審査官引用の特許を対象としている。

Fooust-Jでは日本特許を対象に分析でき、他にも日本特許を対象とした試みがなされている。これらは、特許庁から販売されている「整理標準化データ」から抽出した引用情報を使っており、特許登録前の拒絶引用なども利用することができる。しかし、審査引用が付く前や審査請求されない出願では利用することができない。

引用情報は以上のような分析的利用の他に引用特許を参照するだけのサービスもある。この仕組みは日米特許とも多くのデータベースがもっているが、問題はいかに簡単に利用できるかという点である。最も簡単で便利な利用のできるものの1例として、日立製作所の Shareresearch では抄録と同じ画面に引用特許番号や被引用特許番号が表示される。それらをクリックすることでその(被)引用特許の抄録を参照でき、さらにその引用特許番号をクリックしていわゆる孫引きを次々と展開できるようになっている。これらはキーワードやIPCなどによる検索・調査が基本のシステムであり追加料金なしで気軽に使えるが、ほとんどのシステムは審査官引用を使っている。なお、リコーテクノロジー社の RFPWAY では引用、被引用が抄録とは別画面に表示されるが、このシステムは審査官引用と共に発明者引用も使っている。ただし、発明者引用特許の抽出の精度などについては発表されていない。

また、関西学院大学の玉田俊平太助教授により日本特許の発明者引用の特許を 98%以上の再現率で抽出したとの報告<sup>3)</sup>、<sup>4)</sup>があるが詳細は発表されていない。なお、筆者の調査によると上記以外の複数のデータベースにおいて発明者引用の抽出やその利用が検討されているようであるが実態は報告されていない。

### 3. 日本特許における引用の種類

日本特許制度においては長い間特許公報に記載される引用情報は少なかった。しかし、最近「整理標準化データ」により審査経過情報が入手できるようになり、下記に示すような多くの種類の引用特許が利用できるようになった。

- ①発明者が明細書に書いた引用…この発明者引用は自分の発明の技術的な優位性を説明するために発明者が自ら先行技術を紹介するもの。利用するには本稿で後述するような特殊な抽出作業が必要である。出願から数ヶ月前までに公開された特許が引用の対象になる。
- ②特許庁の調査で抽出された引用…全出願の 6.7 割の審査請求された特許の先行技術を特許庁で調査した結果である。2～6件程度付いているものが多い。また出願から2～5年後になって利用できることが多い。整理標準化データとして入手でき、それらを収録している IPDL(特許庁監修の電子図書館)、その他のデータベースで利用できる。本特許の出願後に公開された特許も対象になる。
- ③審査の拒絶理由通知に使われた引用…審査された特許は一度拒絶されることが多いためほとんどの審査請求された特許で利用可能だが、稀に記載のない場合もある。②の引用特許の一部になっていることが多い。権利を制限する方向で使われるので、①とは異なる特許が使われることも多い。
- ④拒絶確定に使われた引用…特許登録されない場合の理由として引用されたものである。③と同じことが多いが新たな特許もありうる。
- ⑤特許公報に「参考情報」として掲載された引用…特許登録された特許の場合である。拒絶引用が出ている場合はその一部が掲載されることが多いが、新たな特許もありうる。
- ⑥その他、無効審判／異議申し立ての理由として遣われた引用特許やそれらの決定時の引用特許もある。

## 4. 発明者引用の特徴

### 4. 1 発明者引用の長所と短所

特許出願に際し発明者が自ら明細書に書く発明者引用は先行技術に対する自分の発明の優位性を説明するために書かれるものであり、他の引用に比し長所がある一方で独特の問題もある。

- (1) 発明者引用は公開公報により全件が見られるというメリットがある。(これ以外の審査による引用特許は、審査請求され、特許庁で調査が行われた後で付くものであり出願の全体ではない。)
- (2) 公開公報により出願から1年半という短期間で見る事ができる。審査引用の多くは出願から2年～5年も後で見ることができない。
- (3) 自分の特許の優位性を説明するために技術的に広い範囲の特許を引用する場合があります、関連特許を広く探す立場としては長所といえる。しかし、それらは対象としている特許の技術テーマと関係の薄い特許も含まれることもある。このようなことを承知した上で上手に使う工夫が必要である。
- (4) 発明者引用が1件も無い特許がある一方で数百件も記載されている特許もある。しかし、引用がまったくない特許でも他の特許によって引用されること(被引用)は他と同じように期待でき、他の特許とのリンクがなされる。また、引用の多すぎる特許も他の特許から引用される場合は平均的な引用になる。したがって、引用の偏りは、メリットを生かす工夫をすれば克服できる問題であろう。
- (5) 発明者引用は発明者が記載するものであるために記載フォーマットが一定していない。このため、発明者引用の特許番号の抽出には工夫を要する。

### 4. 2 多様な記載フォーマット

筆者は株式会社レイテックの依頼により、明細書中の発明者引用特許を高精度に抽出するシステムを開発した。この開発に先立って筆者は明細書中の多くの発明者引用特許を分析し、特開平5-123号のような典型的なフォーマットの他に多様な記載が混在していることを明らかにし、それらを4つのタイプに分類した。

I タイプ; 標準からややずれた記載法であるが容易に判断できるもの

- ・特許公開平成5-123号
- ・公開特許公報平成第5-123号
- ・特再平5-800123号 ※再公表公報のPATOLISでの表記法
- ・平成5年に公開された特許第123号
- ・特開平5-123号、同5-124号

II タイプ; 元号や特許種別の記載が漏れているが番号が正しいとすれば論理的に判断できるもの

- ・特開60-123号 ※元号が記載されていないが特開昭と判断できる
- ・特平9-123号 ※特公平9はありえないので特開平と判断できる
- ・特許公開9-123号 ←特公平9はありえないので特開平と判断できる

実際には、上記の一部変形、例えばハイフンが長音記号や特殊文字になっているもの、句読点がカンマになっているものや空白、ドット(中黒)になっているもの、など様々のものがある。また、実用新案や公表、再公表、公告、登録の各公報の場合は、さらに独特の記載もある。コンピュータプログラムによる自動抽出は上記のようなロジックの明瞭な記載に対しては対応できる。しかし、下記のタイプとして示すような場合などは自動抽出が困難、または不可能である。

### Ⅲタイプ;論理的に判断不可能のケース

- ・特平5-123号 ※公開、公告の区別ができない
- ・特公15-123号 ※昭和か平成か判断できない
- ・公開特許1456789号 ※記載が間違っている模様

### Ⅳタイプ;前後の特許番号と干渉を起こして判断し難いケース

- ・平成4-123号の公開特許公報は平成5-678号の公告特許公報と・・・ ※後者は公開公報か公告公報か分かり難い

自動抽出率を計算するに当たってはⅠタイプとⅡタイプは当然母数に入れるべきであり、それにⅣタイプも努力すれば自動抽出できるので母数に入れた。しかし、Ⅲタイプは特許番号としての記載に不備があるので母数に入れるべきではないと考えた。この結果、筆者の開発したシステムでは日本特許を対象に99.99%以上の自動抽出を実現した。さらに、特許番号かも知れないが自動抽出できなかった番号もその旨記録することができ、手作業により確認して追加することができる。

なお、2002年ごろの出願から本文中に【特許文献1】などとして発明者引用が記載されるようになったが、そこに書かれているよりも多くの発明者引用が明細書中に書かれることがある。

## 5. 発明者引用の実際

図1は、1993年から2006年までの公開公報の発明者引用特許の平均件数と発明者引用特許が0件の特許件数である。発明者引用の平均件数は年々増加し、発明者引用の全く無い特許は減少していることがわかる。なお、2004年の発明者引用件数が飛びぬけて多くなっているが、これは特許庁の指導による出願時の情報公開の義務化と関係するものと推察される。

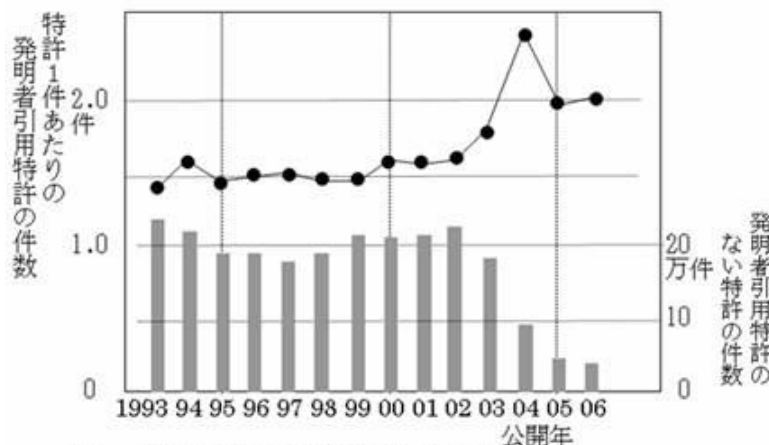


図1 発明者引用の平均件数と全くない特許

図2は、1993年と2006年の発明者引用件数ごとに特許件数を集計した図である。1993年には発明者引用が0件の特許が約22万件(63%)もあったのが、2006年には約5万件(約14%)と大幅に減少している。逆に、1件以上引用している特許は、2006年は大幅に増加している。

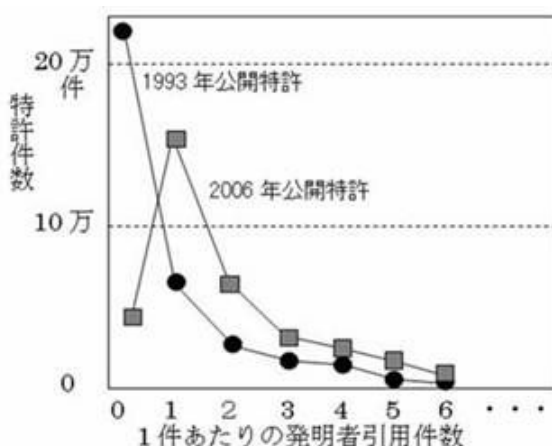


図2 発明者引用件数ごとの特許件数

表1は、2005年公開特許を対象とし、出願国や優先権主張国ごとに見た発明者引用特許の国および特許種別ごとの件数である。優先権主張国が日本、米国、ドイツにおいては自国特許を引用するものが最も多いが、その他のほとんどの国では米国特許を引用するものが最も多くなっていることが分る。(優先権主張国がスウェーデンの場合は米国特許よりWOが若干多い。)

表1 2005年公開特許の発明者引用特許の国・種別内訳

引用特許の種別	合計	元の特許の優先権主張国													
		JP	US	EP	DE	GB	FR	NL	SE	CH	CA	KR	TW	CN	etc
日本特開	543588	540182	1047	86	218	54	129	5	9	3	3	1666	121	29	54
日本再公表	43	42	0	0	0	0	0	0	0	0	0	1	0	0	0
日本特許	50381	49955	180	18	15	11	50	0	0	0	0	50	10	7	85
日本実開	13997	13985	5	0	7	1	5	0	0	0	0	12	1	0	1
日本実用新案	5492	5472	7	0	1	0	0	0	0	0	0	11	0	0	1
日本特願	12751	12487	127	9	9	4	24	0	0	0	0	70	7	1	13
日本実願	406	404	1	0	0	0	0	0	0	0	0	0	0	1	0
US公開	1550	512	667	75	63	36	24	1	1	3	0	109	29	4	16
US登録	44505	19276	17650	1310	1342	332	1431	38	57	45	41	1907	334	32	860
US出願	2607	85	2381	34	22	8	55	0	2	0	0	17	8	0	15
US意匠	27	7	14	1	0	0	0	0	0	0	0	1	3	1	0
EP公開	7194	2237	754	880	1582	193	1121	15	26	70	11	62	8	1	234
EP出願	2	1	0	0	0	0	0	1	0	0	0	0	0	0	0
EP登録	535	22	92	135	39	13	214	3	0	0	0	3	1	0	13
WO公開	11289	7707	1278	527	670	203	433	8	67	21	6	112	9	3	225
WO出願	78	11	37	5	12	2	8	1	0	0	0	0	0	0	2
ドイツ	4276	774	99	274	2834	5	217	2	5	24	0	2	5	0	35
英国	1347	816	144	41	65	103	110	4	9	3	2	6	3	0	41
フランス	1264	235	33	33	20	0	928	1	1	4	0	0	1	3	7
オランダ	18	4	3	1	2	0	0	8	0	0	0	0	0	0	0
スウェーデン	48	5	22	0	2	0	0	0	19	0	0	0	0	0	0
スイス	94	31	4	27	12	0	2	0	0	16	0	0	0	0	2
カナダ	399	239	55	26	49	3	11	0	0	0	13	0	0	0	3
韓国	906	72	3	0	0	0	1	0	0	0	0	827	0	2	1
台湾	238	104	4	0	1	0	0	0	0	0	0	0	128	3	0
中国	208	159	4	2	2	0	4	0	4	0	0	4	4	17	8
その他	1151	869	78	26	23	1	31	0	1	0	0	38	3	1	80
(特許件数)	354900	334474	8868	1108	2390	390	941	52	68	68	20	5079	876	97	468

発明者引用特許件数が極端に多い特許(以下、スーパー引用特許という)が稀にある。表2は、2005年公開特許を対象に100件以上の発明者引用特許のあるスーパー引用特許の件数である。このようなスーパー引用特許は従来技術とか構成部

品や材料においても引用特許をあげて説明しているもので親切といえる半面、対象特許の主題と関連の少ない特許も含まれているので利用に際しては注意が必要である。なお、スーパー引用特許は、特定の出願人に偏っている傾向がある。

表2 スーパー引用特許  
対象：2005年公開特許

引用特許件数	特許件数	引用特許件数	特許件数	引用特許件数	特許件数
100	2	122	1	177	1
101	3	123	1	193	1
102	1	125	1	194	1
104	5	126	1	198	1
105	1	131	1	199	1
106	5	132	2	203	2
107	5	133	1	215	1
108	2	134	1	221	1
109	5	135	1	224	1
110	2	136	1	234	1
111	1	140	2	249	1
113	1	141	1	250	3
114	1	142	3	252	1
116	3	144	1	253	1
117	1	149	1	271	1
118	1	151	1	282	1
119	2	159	1	327	1
120	4	169	1		

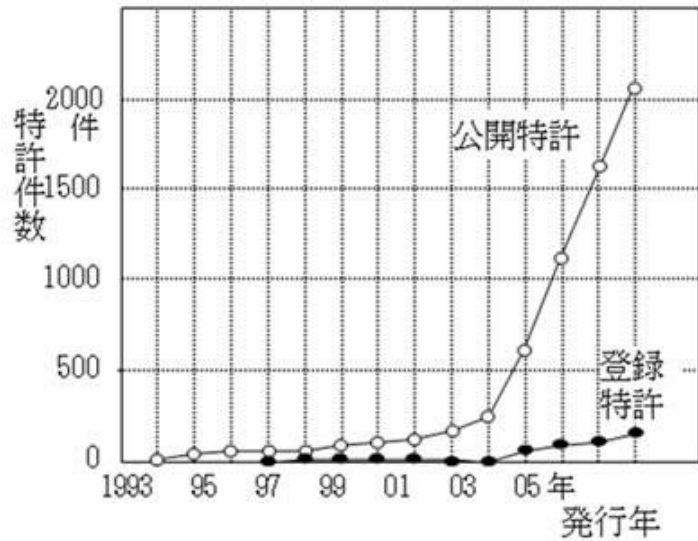


図3 ナノテク関連の特許件数

## 6. ナノテク技術での発明者引用特許

ナノテクノロジー関連特許を題材に発明者引用特許をさらに詳細に検討した。図3は該当する公開特許、公告・登録特許の発行年別の件数である。このように、急増している技術分野では登録公報は数が少なく、引用関係を知ることが大きく制限される。

関連する公開特許 6222 件による発明者引用延べ件数は 32300 件であり、平均 52 件であった。また、公告・登録特許 546 件による発明者引用延べ件数は 2780 件であり、その審査引用延べ件数は 1911 件であった。公告・登録特許の審査引用に対して公開特許の発明者引用で知りうる関連特許は、 $32300/1911=17$  倍もの多くなる。

発明者引用の典型的なものとして下記4例を示す。

- (1)発明者引用特許のみ…例;特開 2002-269757; 合計 11 件の発明者引用がある。しかし、審査請求されていないので審査引用はない。公開特許に比べて登録特許が極端に少ないのでこのようなケースは大変多い。
- (2)発明者引用特許と審査引用特許がある…例;特開 2003-19427; 8件の発明者引用があり公開日の平成 15 年 1 月 21 日以降、利用可能である。本件は登録になっており2件の審査引用があり登録公報発行日(平成 18 年 12 月 13 日)以降、利用可能である。登録公報の多くはこのようなケースである。
- (3)審査引用特許のみある…例;特許第 3081842 号; 合計3件の審査引用が拒絶引例として平成 12 年 4 月 11 日に出ており、これから 2、3ヶ月後には、整理標準化データとして利用できる。また、公報発行日の平成 12 年 8 月 28 日以降は同じ特許が引用特許として公報に掲載されているので利用できる。これに対し、公開公報での発明者引用はない。このような例も時々ある。

(4)登録特許だが審査引用はなく発明者引用のみある…例;第 2765543 号;審査引用は 1 件もない。これに対し、その公開公報である特開平 9-175870 は2件の発明者引用特許があり、その公開日(平成 9 年 7 月 8 日)以降利用可能である。稀にこのようなケースもある。

## 7. 終わりに

発明者引用は網羅性、速報性などのメリットがあるにもかかわらず、その詳細は分からなかった。本報告により発明者引用の実態が明らかになり、体系的な利用に資する情報が提供された。今後は発明者引用データを活用するシステムの開発が望まれる。

(基礎データをご提供いただいた株式会社レイテックの出口隆信社長に感謝申し上げます。)

## 参考文献

- 1)特許のはかり方、1997.11.30、翻訳;松山裕二(ゼファー株式会社一現、ゼファービヨンド株式会社)
- 2)日本特許協会、引用特許シンポジウム予行集、平成 16 年 7 月 2 日
- 3)特許による知識の移転はどの分野で特に有効と考えられるか? JAPIO 創立 20 周年記念誌、P.135、2005.10.1、玉田俊平太(関西学院大学助教授)
- 4)企業における基礎研究とイノベーション、—日本の大手電機メーカーの特許性向と科学依存度—、2005 年 1 月、鈴木潤(財団法人未来工学研究 R&D 戦略研究センター)、玉田俊平太、他

英文タイトル Inventor citation patents and development of its extraction system

Masamichi MUGURUMA(Muguruma Professional Engineer Office,4-37-6,Nishinarusawa-cho,Hitachi-shi,Ibaraki-ken,316-0032 JAPAN)

Abstract: Some kinds and its features of citation documents in Japanese patent system were described. The various formats of the Cited Patent by Inventor (CPI) written by inventors in Japanese patent specifications were analyzed. A high precision CPI extracting system from the patent specification was developed by the author. The real status of the CPI from 1993 was uncovered by the article. In the nanotechnology field as example, it was showed that the CPI was more useful tool than cited patent by examiner.

Keywords: citation patent/cited patent by inventor/citation extraction/cited patent by examiner/nanotechnology