

特許情報検索と解析の将来展望

個人サーチャーの記憶限界を超える FTO 調査

一般財団法人日本特許情報機構 特許情報研究所客員研究員 桐山 勉

PROFILE

2006年4月末から(財)日本特許情報機構の客員研究員を務める。2005年からINFOSTA-SIG-パテントドキュメンテーション部会のコアパースンとして活躍中。他にも特許情報検索競技大会の副委員長など、各種組織と関係を持つ特許情報調査のマルチプレイヤー。



1 はじめに

2009年4月に岐阜県大垣市の片田舎に転居後も、ライフワークとして「特許情報を通じてイノベーションに資するR&D効率アップ」という目標に向かってNPO的な寄与ができないか模索中である。

高齢になるにつれて記憶力と集中力が若い頃に比べて数分の一に衰えていくのを毎年痛感している。人間が記憶力に依存しながら特許情報を精査・判断できるのは200件~250件と言うのが持論である。一方、特許情報調査の世界では扱う特許情報の数量が益々膨大になり、数万件を扱う技術経営(MOT: Management of Technology)の実務において、イノベーションに資する特許情報解析が履行ベースで年々難しくなっている。

筆者は独立行政法人工業所有権情報・研修館が主催開催している特許情報検索競技大会の副委員長も務めているので検索競技大会の制限時間と同じ5時間で如何に集中力と記憶力と経験知 Skill を使って特許情報の調査を人的能力中心で履行するかに大変興味を覚えている。

その一方で、記憶力をはるかに超えた数万件の特許情報のFTO(Freedom to Operate)調査を一カ月ぐらいの期間で解析ツールを駆使して如何に実践するのかにも同時に興味を持っている。

この両方の特許情報調査解析を習得する為に、(1)定期的な定点観測¹⁾と称してこの業界の情報を積極的に収集する努力をしている。(2)知ったことを更に調べる

ために、自分なりに調査できる商用データベースを複数使える状態を保つ努力と工夫をしている。(3)更に、積極的に時間を割き、できるだけ国際会議(例えば、IPI ConfEx2010²⁾、米国PIUG Annual Conference 2010³⁾、PIAC in China 2010⁴⁾等)に参加するように老骨に鞭を打っている。国際会議の後でダウンロードできる発表パワーポイント資料とPIUG登録会員の為にmail配信されるwiki-piug資料は、大変勉強になる。この一年間に筆者が何度も読み返す文献・記事を参考までに5つ挙げる。

- 1) Palmisano Report (Innovate America)⁵⁾
- 2) イノベーションのための特許情報の活用(鶴見教授)⁶⁾
- 3) テキストマイニングと統計解析言語Rによる特許情報の可視化(安藤氏)⁷⁾
- 4) How text mining is changing the way you search (Wolfgang Thielemann, IPI-ConfEx2010)⁸⁾
- 5) Toyota settles with Paice on hybrid car litigation.(PIUG membership e-mail delivery)⁹⁾

これらの5つの記事に共通する考えは、筆者の解釈によれば、次のように映る。知財管理の要諦は特許情報である。高度な特許情報解析こそがイノベーションに資する実務である。特許情報解析の生産性を向上させるためには、テキストマイニングによる俯瞰可視化と通常のマップ・ツールによる情報加工は避けて通れない。

2 FTO調査を履行するには何が必要か

数万件の大規模集合群の特許情報の中から価値ある中規模集合群（約 2,000 ～ 3,000 件）の特許情報を選び出し、最後に重要な価値を持つ小規模集合群（約 200 件～ 350 件）に絞り込むフィルター（漏斗）プロセスが必要になると筆者はモデル図で考えている。

つまり、FTO 調査を履行するためには実務として三段階で選択と抽出ができる情報ハンドリング技術の能力が必要である。筆者は、IP Information Scientist¹⁰⁾ の能力として 6 つの能力¹¹⁾ を主唱している。中でも技術理解力、検索技術力、知財実務力、IT 力、情報解析力、基礎力（コミュニケーション力を含む）の順序で重要と考えている。そして、FTO 調査ができる人がスーパーサーチャー¹²⁾ と信じている。FTO 調査

には記憶力をはるかに超える膨大な特許情報を自由自在に扱う情報加工力が必要になる。

3 トヨタ社とPaice社の訴訟に使われた解析とは

前述の第 5 文献・記事は、OPTIMICE 社の Mike Lloyd が配信した記事であり、「トヨタは、Paice 社とハイブリッド自動車の特許係争を和解し、その際に、オーストラリアの Griffith Hack 特許法律事務所が OPTIMICE¹³⁾ 社より導入して活用した NWA 解析 (Network Patent Analytics) が調査の決め手になった」という内容であり、筆者は釘付けになった。

その理由は、図 1 に示した大規模集合群から中規模集合群へ、最後には小規模集合群へ、どの様に実務として絞り込みを実施したのか知りたかった、真似をし

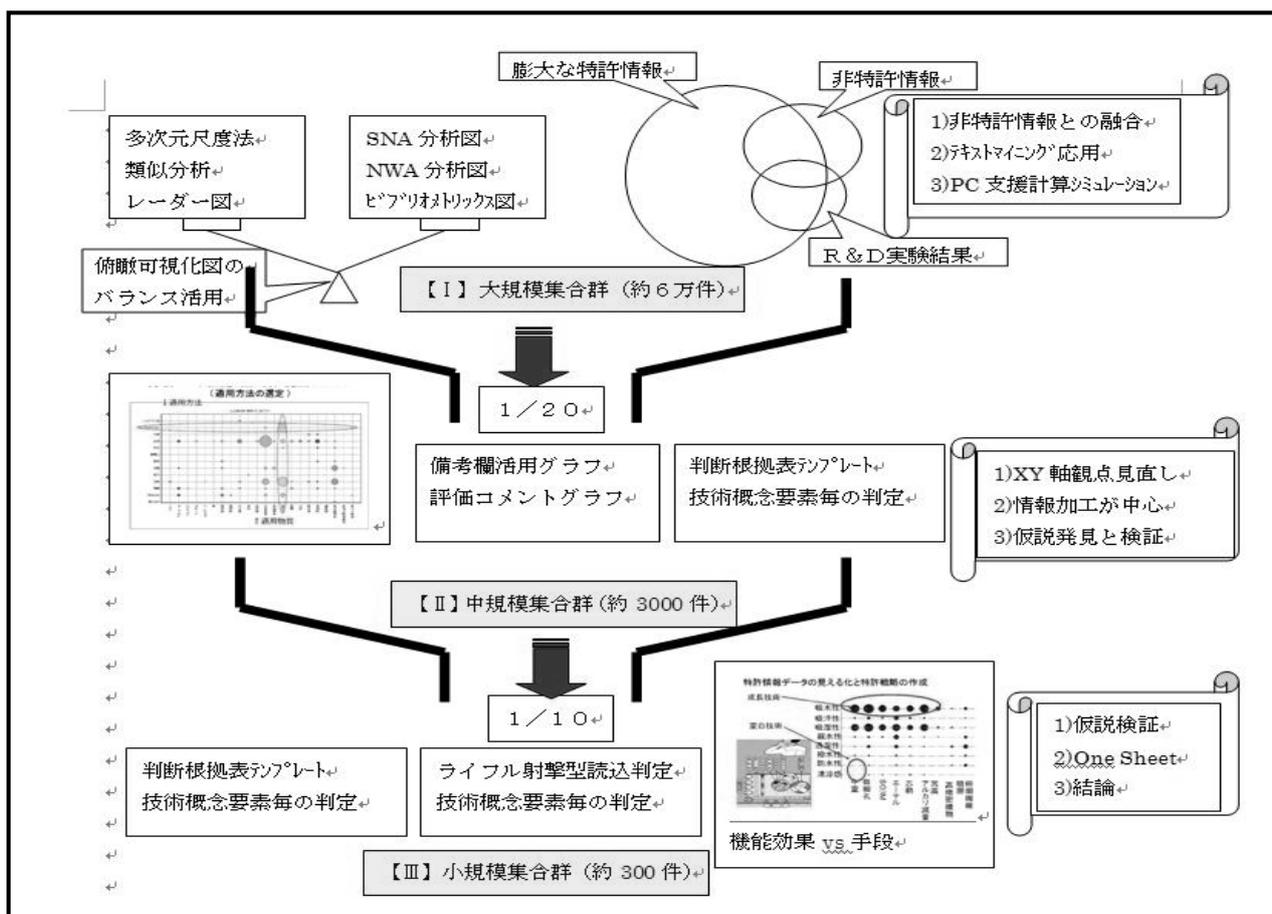


図 1 抽出選択の三段階プロセスのモデル図

たかったからである。PIUGの記事によれば、6万件以上の特許が両社の特許訴訟に絡んでいるとのこと。特に、訴訟の的は、Paice社のUS特許5,343,970とトヨタ社のUS特許5,841,201とトヨタ社のUS特許5,839,533とであることをNWA解析でつきとめたと述べている。約6万件の特許情報を通常の解析手法で履行するには一人の個人サーチャーの記憶力と集中力の限界をはるかに超えていると考える。そのため、このような複雑な係争に絡む特許情報の解析にはIP Information Scientist(知財情報専門科学者)またはスーパーサーチャーが欠かせないと筆者は考えている。特に複数のスーパーサーチャーがプロジェクト・チームを組まないと履行できないほど、最近の特許解析は複雑で高度でツールに慣れることが必須だと信じ込むように筆者はなった。筆者はINFOPRO2008にて、「R&D効率アップの特許情報活用—書誌ブリーアン系・俯瞰可視化系・通常マップ系の連携活用研究—」を発表した。いち早く、パテントロール問題の被害を予防するためにも、日本の企業の特許情報解析力を向上させることが急務であると感じ始めていた。

OPTIMICE社のHP資料をインターネットで調べてみると、NWA解析には俯瞰可視化として図2および図3の手法を使っていることを示唆している。

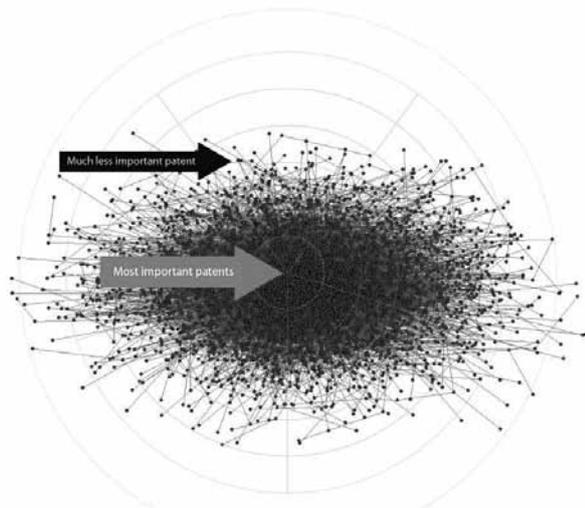


図2 大規模集合のNWA可視化図(OPTIMICE社のHPより転載)

図2は、レーダー図であり、注目特許が中心に位

置するように配置されている。SO-TI社のχLUS-White図¹⁴⁾に酷似している。一方、図3は、World Patent Information専門誌に紹介されたSocial network Analysis¹⁵⁾で得られる図と似ている。

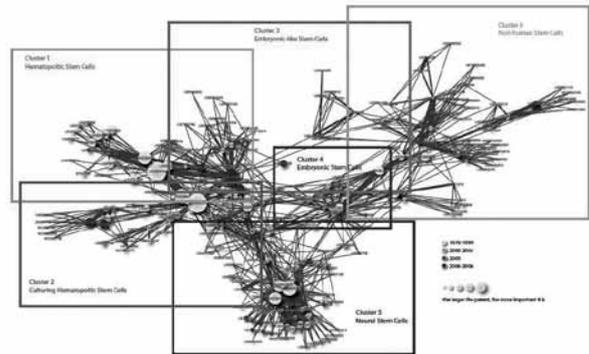


図3 大規模集合のNWA可視化図(OPTIMICE社のHPより転載)

最近では、NRIサイバーパテント社のTrueTeller¹⁶⁾で可能なNWA分析と似た機能で作成されるものと判断する。INFOSTAより最近発行された新訂情報検索の知識と技術の本の中で、ビブリオメトリックス¹⁷⁾と紹介されている可視化図と酷似している。だから、日本においても、SO-TI社のχLUS-WhiteとNRIサイバーパテント社のTrueTellerのNWA分析を使い、それらの書誌データを使い通常の特許マップ・ツールを活用すれば、約6万件→約3000件→約300件の抽出選択の三段階プロセスは実行ベースで追試も実践もできると信じている。この三段階プロセスを一サイクルすれば、品質的に80点以上のFTO調査ができると信じている。更に、三段階プロセスの見直しをもう一回サイクル分を行えば90点以上の品質のFTO調査に改善できると信じている。更に最後にもう一回サイクル分を行えば95~99点以上の品質のFTO調査に最終的にできると信じている。

4 通常の古典的な特許情報解析の限界

それでは、高度な特許情報解析と対局に位置する通常の古典的な特許情報のやり方とは、どのようなモノで

あろうか。通常の解析手法とは、表計算ソフト、例えば、EXCELのマトリックス表に縦方向に約6万件の特許番号リストを作成し、横方向に解析しようとする特定技術の必須技術要素概念と必要技術概念要素を細かく分けて、明細書を精査して、それぞれに細分化された技術概念要素ごとに「A:該当」または「B:参考関連」または「C:無関係」と判断する方法が最も基礎的なやり方である。筆者は、このやり方をライフル射撃型の特許情報解析と呼んでいる。実務的には、小規模集合群と中規模集合群の特許情報解析の明細書読込に活用する常套手段である。しかし、小規模集合群の特許情報解析に使うEXCEL表の常套手段を約6万件までに拡張するには時間が掛かり過ぎ、現実的ではない。 $60,000/500 = 120$ なので120倍の長時間の努力を継続すれば計算上は可能であるが、個人の人間の繰り返し実務の限界を超えている。

更に、少し工夫をして市販の特許マップ・ソフト（例えば、レイテック社のPat-Listソフト、インパテック社のEXZソフト、日本パテントデータサービス社のJP-MAPソフトなど）を活用しながら明細書を読破するやり方がある。但し、特許マップ・ソフトと特許検索システムの連携機能が必要である。連携していないと、特許番号などの再入力などに時間と精力が使い果たし苦労が絶えない解析作業となり、筆者はそのような非効率な特許情報解析は薦められない。この方法でも約6万件の特許情報解析は、個人の人間の繰り返し実務の限界を超えている。 $60,000/2,000 = 30$ なので30倍の努力で計算上は可能に見えるが、個人の繰り返し実務の限界を超えている。

5 ビッグバンで出現した特許情報解析法とは

IP Information Scientistが複数でプロジェクト・チームを作り、SO-TI社の χ LUS-WhiteとNRI社サイバーパテントのTrueTellerのNWA分析を使えるのなら、それらの書誌データを使い通常の特許マップ・

ツールを活用できるのなら、高度な特許情報解析が現実的に日本でも可能と信じる。筆者は、3万5千件の特許情報の書誌データを解析した（特許庁のHPに掲載された）特許マップを二テーマ、合計二年間の経験がある。数千件の特許情報を通常の特許マップ・ツールを活用して特許情報解析をする経験も実務として複数回、経験することができた。SO-TI社の χ LUS-WhiteとNRIサイバーパテント社のTrueTellerも使った経験がある。しかし、筆者には、約6万件のFTO調査をした経験がない。約6万件のPATOLIS抄録・クレーム・書誌情報と、SO-TI社の χ LUS-WhiteとNRIサイバーパテント社のTrueTellerを駆使して、抽出選択の三段階プロセスを用いて約2カ月間で特殊な技術テーマにおけるFTO調査を三人の有志で一度、やってみる経験も積みたい気がする。

ビッグバンで出現した特許情報解析法とは、約6万件という膨大な特許情報でも扱うことができるSocial network Analysisとか、ビブリオメトリクス分析とか、NWA解析(Network Patent Analytics)とか、 χ LUSのような多次元尺度法類似解析などを意味することがここに来て漸く具体的に理解できた気がする。欧州特許庁の副長官Curt Edfjall氏¹⁸⁾が2008年にWorld Patent Information専門誌に掲載した時点では、テキストマイニング分析を応用した技術の何かが出現するのだろうかと思った。我々は特許情報解析手法がまだまだ発展する過渡期に生きている。

6 Wolfgang Thielemann氏の考え方

前述の第4番目の発表記事「How text mining is changing the way you search (Wolfgang Thielemann) は、IPI-ConfEx2010でのTutorialで説明されたことを後日にダウンロードできた発表資料⁸⁾で初めて知った。彼は、4つのサーチ方法(Boolean Search, Explore & narrow down, Semantic similarity, Boolean searching & Text minig)を主唱

しており、筆者も類似した考えを持っていると自覚し、彼の発表資料を何度も読み返した次第である。大変参考になると考えるので、IPI-ConfEx2010 から転載する。

7 最後に

結論を述べると、知財管理の要諦は特許情報である。高度な特許情報解析こそがイノベーションに資する実務である。特許情報解析の生産性を向上させるためには、テキストマイニングによる俯瞰可視化と通常マップ・ツールの情報加工の併用は避けて通れない。

「世界を知り、世界と共に生きる。」をライフワークの基本方針に加えたいと思った。2010年3月にIPI-ConfEx2010に参加し、5月にPIUG Annual Conference2010に参加した。今迄、雲の上の専門家と思っていたStephen Adams氏、Lucy Akers氏らと短い時間であるが談笑ができた。米国にはPIUGがあり、欧州にはCEPIUGが在ること知った。日本にはPIUGがない。この状態は、「世界を知り、世界と共に生きる」という基本方針から外れている気がしている。日本だけが世界から取り残されないために、日本のサーチャー同士がNetworkを構築し連携する必要があると痛感した1年間である。

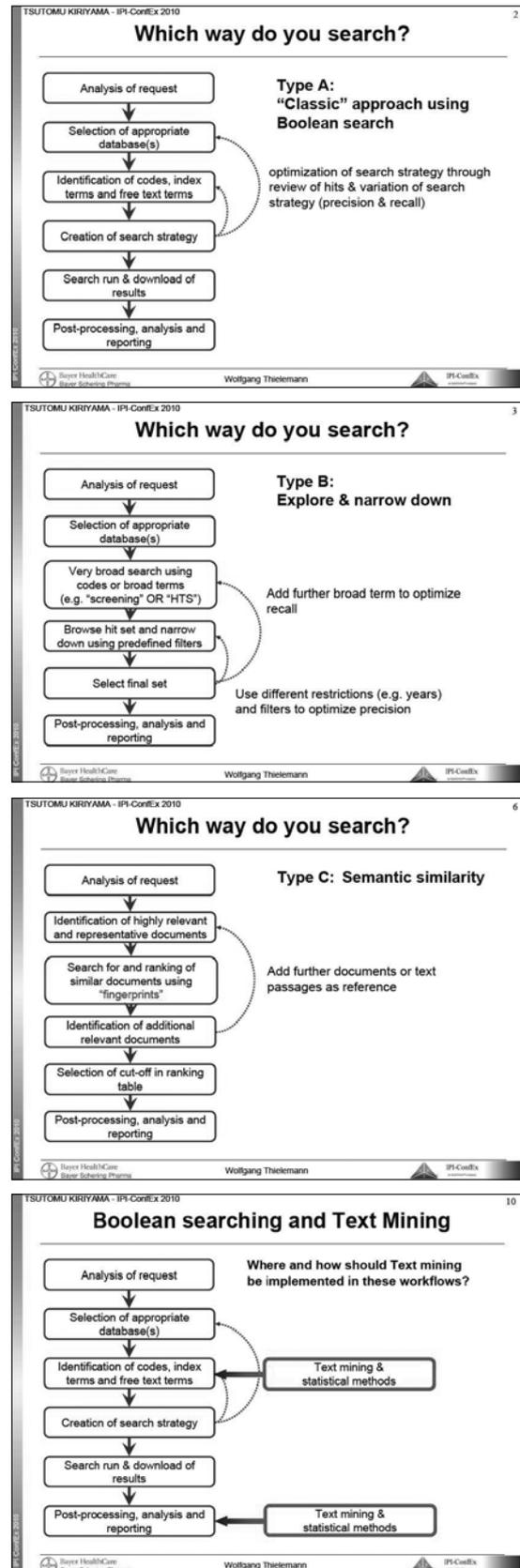


図4 Wolfgang Thielemann 氏の解説

[参考文献]

- 1) 筆者の情報収集の定点観測：Japio 2009 YEARBOOK, p209.
- 2) IPI ConfEX2010: <http://www.ipi-confex.com/>
- 3) PIUG Annual Conference2010 :
<http://www.piug.org/2010/an10meet.php>
- 4) PIAC in Chna2010 :
http://www.cnipr.com/piac/eng_program.html
- 5) Innovate America , December 2004.
http://www.compete.org/pdf/NII_Final_Report.pdf
- 6) 鶴見隆、イノベーションのための特許情報の活用、知財管理 ,Vol.60(3),p375-392(2010)
- 7) 安藤俊幸、テキストマイニングと統計解析言語Rによる特許情報の可視化、情報管理 Vol. 52 (2009) , No. 1 p.20-31
- 8) Wolfgang Thielemann,IPI-ConfEx2010、Tutorial, Lisbon, March 9th,2010.
- 9) Toyota' s litigation with Paice: PIUG 登録メカニズム専用配信 wiki 資料 2010/07/28/
- 10) IP Information Scientist: 関西特許情報センター振興会、機関誌 No.22,p3-10(2009)
- 11) サーチャーの6つの能力：桐山勉、Japio2007YEARBOOK,p108(2007.11月)
- 12) イノベーションに寄与する特許情報検索業務者の育成 “スーパーサーチャー” を創る；渋谷善弘、Japio 2009 YEARBOOK 、 p100-103、
- 13) OPTIMICE 社：Network Maps.
<http://www.optimice.com.au/maps.php>
- 14) SO-TI 社（創知）、XLUS – white マップ
<http://so-ti.com/service/xlus.html>
- 15) Visualizing patent statistics by means of social network analysis tools.: WPI Vol.30, (2008) p115-131
- 16) NRI サイバーパテント：TRUE TELLER
<http://www.patent.ne.jp/service/macro/index.html>
- 17) ビブリオメトリックス：新訂情報検索の知識と技術、(社)情報科学技術協会発行、2010年5月第3版、p203
- 18) The future of European patent information. Curt Edfjall. WPI Vol30 (2008) 135-138.

