

特許情報検索と解析の 将来展望

検索システムはどこまで改善が必要か

株式会社帝人知的財産センター、OB嘱託
桐山 勉

PROFILE

約20年間の研究開発の後、1994年から知財部門で特許・文献調査に従事し、帝人グループの調査とシステム構築を担当した。2006年に定年退職しその後はOB嘱託として勤務すると共に、INFOSTA-SIG-パテントドキュメンテーション部会のキーパーソンを継続中。

✉ | t.kiriyama@teijin.co.jp

☎ | 03-3506-4057

1 はじめに

昨年の2007YEARBOOKにおいて、特許情報解析の将来展望について述べた。その後、1年が経過し、知財情報を各企業の研究開発（R&D）の効率化に如何にリンクさせ、知財の付加価値を如何につけ開発製品を如何に守るかが、研究技術者の日常的な認識になりつつある。しかし、その意識の温度差はますます広がり、意識の高い「一部のグループ」と相変わらず研究活動の多忙性に埋没し知財情報の整理と解析に時間が十分に割けない「大多数のグループ」に二極化している様な気がしてならない。各企業では定額制の特許検索システム（ASPサービス）を如何に定着化させ、如何にバージョンアップさせ、社内に教育普及させるかがポイントになりつつある。内外の特許情報システムに関する情報を集めながら、欧米のIP Professional Information Scientist（以降、知財プロフェッショナル情報科学者と略す）が何を考えているのか考察する。皆様のご意見を頂く機会

にねれば幸いである。

2 FTOとFulgurationとは何か

筆者は欧米の知財プロフェッショナル情報科学者が何を考えているのかその動向を把握するために、下記の4団体の年次大会の発表に注目している。

- (1) IPI-ConfEX¹⁾（通常毎年3月頃、欧州で開催）
- (2) 米国PIUG年次大会²⁾（通常毎年5月頃開催）
- (3) EPOのEast meets West³⁾（毎年4月頃開催）
- (4) EPO-PIC⁴⁾（通常毎年秋、欧州で開催）

これらの国際会議はインターネットで参加者を応募しており、日本の特許情報担当者が強い参加希望を持ち費用を負担すれば自由に参加できるものである。筆者は過去に上記の（3）と（4）には実際に参加した経験がある。【図1】に（3）の写真を示す。

一方、これらの動向を把握するために、日本から実際



図1 East meets West 2008の会場風景と筆者（中）

に上記の国際会議に参加された人とのhuman networkを財産とし、その人達からの情報収集も大切にしている。

更に、専門雑誌としてELSEVIER社発行のWORLD PATENT INFORMATION誌⁵⁾ (以降、WPI誌と呼ぶ)を重要な情報源としている。筆者はINFOSTA-SIG-パテントドクメンテーション部会⁶⁾のコアパースンを務め、メンバーが興味を持った注目記事を自己申告制の輪講会にて相互に紹介して貰っている。WPI誌には前述の4団体で最初に発表されたものを投稿用に補充された記事が時々掲載され、情報源として役立つ。

IPI-ConfEXの会議ミニ報告がWPI誌⁷⁾に報告され、欧米の知財プロフェッショナル情報科学者がFTQに大変興味を示していることが掲載されている。FTQとは、“Freedom to operate” searchの略号であり、企業が自社の新製品を他社の特許に抵触することなく自由に生産できる知的財産権 (IPR) をどの程度保有しているかを調査することを意味している。従来の表現で言えば、権利調査とか抵触性調査とか侵害性調査と言われていた調査法である。知財プロフェッショナル情報科学者が中心となり、SDI情報と通常の調査結果とファミリー特許を中心とした引用被引用検索も組み合わせ、更に、社内の研究者と専門家のコメントも組み合わせた調査が有効と記載されている。検索システムにはその様なコメント欄を標準機能として有することが重要であると言う発表もされている。特に、近年は特許情報が複雑になり、知財プロフェッショナル情報科学者でない研究者の調査では間違っただ判断に陥り易いと警告している。

自社の出願特許群を番号入力したら、その周辺に関係する他者の特許群がレーダ図の如く、出願人別にカラーで全体俯瞰的に表示できるシステムが期待されている。この機能に近いのが、NRIサイバーパテントのTrueTellerと、創知のxLUSと筆者は考えている。しかし、この二つのシステムも将来はさらに改善されると

信じている。

一方、Fulgurationとは電光石火の如く出現するビッグ・バンと言われ、欧米で注目されている。このFulgurationの一つの候補は、サーチエンジンと機械翻訳ソフトウェアの組合せであり、もう一つの候補はWeb2.0の機能を組合わせた特許情報の処理であると言われている。色々な新しいIT技術を応用した情報処理技術が生まれて来るので予想が難しい。日本における特許情報検索の各種エンジンもこれらの新しいFulgurationの影響を受けて改善工夫がされるものと筆者は予測する。

3 欧州特許庁Edfjall副長官の指摘⁸⁾

2007年は欧州Patent Documentation Group (以降、欧州PDGと略す)の50周年記念年に当る。欧州PDG⁹⁾と欧州特許庁は永年に亘り協働作業をして来た。欧州特許庁は1978年から活動し2008年は30周年に当たる。2008年5月には米国PIUG年次大会²⁾が20周年記念として盛大に行われた。筆者もPIUGの有料会員になっているのでmembership onlyのサイトにおいて実際に発表されたプレゼンを参照することができる。

WPI誌には欧州特許庁Edfjall副長官の記事が最近多く掲載されている。彼が色々な所で講演していることがわかる。その中で、「世界の特許情報をインターネットと各種のシステムで利用している人数は間違いなく格段に急増しているが、適切な高度な検索をしている専門家は返って少なくなっているのではないかと危惧している。【図2】にEast meets West 2008の懇親会でEdfjall副長官を囲んで撮った写真(筆者は最左)を示す。また、筆者も昨年の特許情報検索競技大会のフィードバック・セミナー¹⁰⁾にて、同じことを切実に感じている。筆者

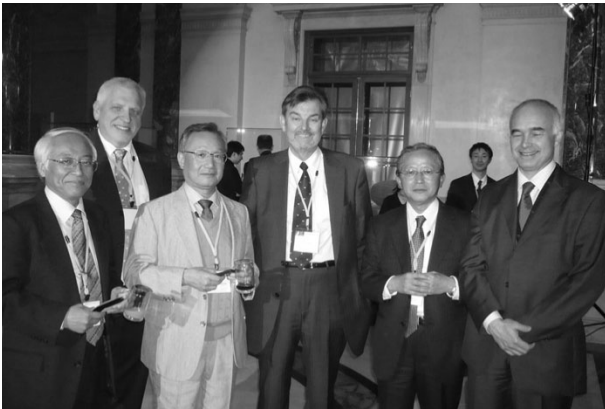


図2 EMW2008にてEdfjall氏（中央）と筆者（最左）



図3 Minoo会長（最右）と筆者（右2番目）

は知財プロフェッショナル情報科学者の役割が見直されると信じている。欧州PDGのMinoo会長らと今年の4月に直に面談して筆者は特にこの意を強くした。Henkel社で撮ったその際の写真を【図3】に示す。

4 特許分類コードは国際的な文化遺産

国際特許分類（IPC）は2006年1月から第8版が全世界共通の特許分類として利用されている。この特許分類も原則として5年毎に版を改定されて来ており、世界的な技術の遺産である。特許分類コードには、国際特許分類（IPC）の他に欧州特許庁分類（ECLA）、米国特許

分類（USC）、日本特許庁分類（FI、Fターム）などがある。これらのメンテナンスに膨大な（人的、システムの、費用的）エネルギーが使われている。

言語により特許明細書のどこに何が記載されているのか理解され易くする為にINIDコードが公的機関WIPOで定義が決められている。その日本語版の解り易いもの¹¹⁾もインターネットで公表されている。特許情報の調査をする際に大変便利で重要な分類定義書である。

しかし、この重要な世界的な遺産である特許分類コードは適正に活用されているのだろうか。特許情報の社内教育を行っている際に痛感する素朴な疑問である。

同様な疑問を欧州庁のPaul Schwander氏は「まだ世の中で見たことが無い」症候群¹²⁾と指摘している。世界の研究者に国際特許分類（IPC）を有効に活用して貰うためには検索システム側で工夫と改善が必要である。2001年に開催されたJPINFO会議（EPO主催、ウィーン）にて筆者は、「IPCは研究者にとり暗号であり、この意味を表示する検索システム側の改善が必須である」とプレゼンし、そのユニークな主唱を欧州の専門家に注目されたことを記憶している。

日本の特許検索システムにおいてこの機能を早くから標準としてサービス提供してきたのは富士ゼロックスのDocuPat検索システムである。その他ではPATOLISのランク機能が特許分類コードの意味解析に極めて便利である。2007年に開催された特許検索競技大会において第1希望の商用検索システムの他に第2の商用検索システムを設けた理由は、特許分類コードの解析を短時間に正確に達成するためにはDocuPatとかPATOLISの分類統計処理と意味表示が必須と考え、PATOLISのランク機能を参加者全員に利用できる環境を準備すべきと考えたからである。最近では、日立情報システムズのSharesearchにもこのランキング機能が【図4】の様に標準装備されたが、意味表示がリンク機能でな

れていない。これらの検索システムには、まだ改善の余地がある。

5 国際特許分類 (IPC) は減びる 巨大爬虫類か

日本における最近の特許検索システムにおけるテキストマイニング機能、クラスタリング機能、多次元尺度法解析機能などの発展ぶりを見ていると、特許分類コードなどは使わなくても明細書のテキスト本文だけを解析すれば十分な特許調査に近い将来に出来る様になるのではないかと錯覚を覚えそうである。しかし、この錯覚の方向の推論は正しいのであろうか。この方向が正しいと信じる人にとっては、特許国際分類 (IPC) は将来使われなくなる可能性が高くなり、丁度、地球上の巨大な爬虫類が減びた様なものに喩えられるのかも知れない。

しかし、ここで特許情報の検索の王道で基本中の基本である「推定再現率と適合率」に注目したい。現時点のテキストマイニング機能、クラスタリング機能、多次元尺度法解析機能などで明細書の全文を対象にしたもので

再現率が50%を超えるものがあるとは筆者は信じられない。概念検索とか類似検索では重み付け検索などと称して特許分類とか出願人などの組合せ検索をしない限り、推定再現率が50%を超えるものは筆者の調査経験から見当たらない。筆者の経験から推定した「推定再現率と適合率」モデル図を【図5】に示す。

インフォプロのベテラン・サーチャーが検索式を作成する場合には、Fugmannの概念の三角形¹³⁾をいつも念頭に置き、表現の多様性を先ず考える。また、特許分類コード類を正確に十分に駆使して検索を作成する。そのため、ベテラン・サーチャーと経験の浅いサーチャーの差が生じるのはこの特許分類コード類を正確に理解し十分に使いこなしているかに掛かっている。このことから容易に理解できることは、現実には特許分類コード類を使わずに推定再現率を80%以上の検索精度を持つ特許調査が出来ないのである。以上から、特許国際分類 (IPC) はそう簡単には使われなくなるとは信じられない。逆に、如何にエンドユーザである研究者に特許分類コードを適切に活用して貰うためには、検索システムをどのように改善設計すべきかが重要である。この機能を

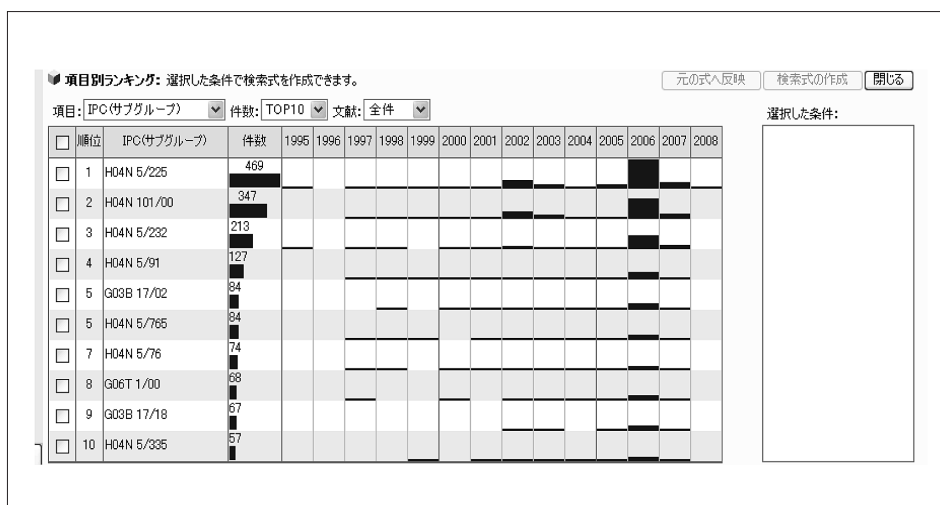


図4 Sharesearchのランキング機能

一番進化させたプロバイダーの検索システムだけが平成22年でも生き残ることが出来ると予想する。特許国際分類は少なくとも今後10年間は減びることがないと断言する。この世界遺産をシステムの的にもっと重用したら良いと考える。

6 R&D効率を向上させるシステムは何か

特許検索システムに今後は次の様な機能が重要視されると筆者が考えているものを列挙する。

6.1 表現の多様性を支援する機能

人間の脳の働き、脳科学の見地、及び発想法とから「表現の多様性」を検索者に容易に気付かせる支援ツールが求められている。「シソーラス辞書の機能を見せる化させた全体俯瞰図」が提供されると便利な支援機能になる。そこで注目しているのが富士通のATMS/Analyzerのスケルトンマップ図と日立製作所のDualNavi図である。検索式を作る前に、例えば概念検索か類似検索の味見検索の小集合から例えばスケルトン

マップ図またはDualNavi図を作成する機能が一般の特許検索システムの標準支援機能として装備されることを提案する。筆者はJPDS主催の特許分析セミナーの講師も担っており、2008年にこのATMS/Analyzerもセミナー演習で利用できる様に企画し、スケルトンマップ図の見せる化により「表現の多様性」を容易に気付かせることができるか実証実験を行い、その効果の確証を感じた。

6.2 特許分類コードの統計と意味表示のリンク

次に希望する必要機能がランキング機能のさらなる進化と改善である。特許分類コードなどの統計ランキングは即座に表示される様になったが、最も重要なその意味がリンクして表示されない。検索者は特許庁IPDLの Patent マッピングガイダンスで特許分類コードを入力して初めて重要な意味が理解できるのが現実である。個々の検索ユーザの貴重な時間を無駄使いしている。必要な情報に到達するアクセス時間をゼロにする点からも是非とも実現させたい支援機能である。

6.3 二画面機能

次に欲しいこの機能は、検索システムの機能よりも利

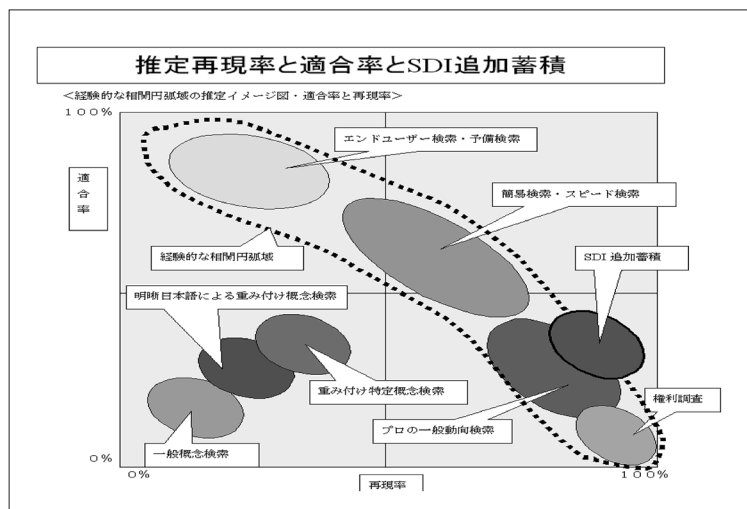


図5 推定再現率と適合率の推定モデル図

用法の問題でもある。特許庁審査官は随分前から特許調査に二画面を利用している。例えば、下記の操作を筆者は想定している。

- (1) 抄録と全文またはレイアウトを同時に参照
- (2) 統計ランキングとその意味を同時に参照
- (3) 特許情報と非特許情報を同時に参照。そして、特許情報にフラグ（しおり）を付け重要視する。
- (4) 報告書と特許検索システムを同時に参照。

筆者は右目の視力が悪いために如何にしてスピードを落とさない様にするか人一倍の努力と工夫が必要になっている。そのため、約2年ほど前から特許調査および分析・解析の際には二画面を使っている。一方でランキング機能を使いながら、他の画面にて特許庁IPDLの Patent マップマップガイダンスをコピー&ペーストにて参照し、ワンクリックでも1秒の操作でも省略化できない工夫している。

6.4 全体像の俯瞰図作成機能

次に欲しい必要機能は特許情報の集合全体の見える化とポジショニングである。筆者はINFOSTA-SIG-パテントドキュメンテーション部会のキーパーソンも務めるために、「書誌ブリーアン系－俯瞰可視化系－マップ系の連携活用研究」¹⁴⁾を部会活動として行っている。創知社のご好意でχLUSのレーダ図を用いて、自分の研究または調査したい案件の位置確認（ポジショニング）を行っている。同じ様なことが、NRIサイバーパテントのTrueTellerでも、富士通のATMS/Analyzerのスケルトン図とアンカー図でも出来る。人間の脳を支援する支援ツールとして小職の経験から、全体俯瞰図作成機能は欲しい機能である。

6.5 反転ハイライト機能

次に欲しい機能は人間の視覚を支援する機能である。黒一色の特許情報は集中して読むのがつらい。筆者は、セミナーの講師をして感じたことがある。同じ特許の明

細書を限られた時間内に正確に読む際に、同じ会場の中で参加者の行動に違いがあることである。例えばとしてA,B,Cの三人の参加者を比較したい。参加者Aは反転ハイライト機能を使っていない。何故使わないのか筆者には解らない。機能があること、その操作を知らないのだろうか。参加者Bは明細書1件ずつ、または抄録1件ずつを反転ハイライト機能を利用しながらチェックしている。参加者Cは一括抄録表示機能を使いながら反転ハイライト機能も使っている。また時々、明細書本文をチェックしたい際にもWindows-XPのマルチウインドウ機能で一括抄録ハイライト窓とハイライト明細書窓を瞬時に切替ながら両方ともハイライトさせながらチェックしている。筆者なら、当然、参加者Cのチェック方法を推奨したい。

6.6 FTO機能

次に欲しい機能が二社または三社の比較機能である。つまり、前述のFTO機能がスピーディーに出来ることである。例えば、インパテック社のテクレスZには二社の比較機能がある。自社の特許群と競合ベンチマーク会社の特許群の比較機能である。また、富士通のATMS/Analyzerのアンカー図により3箇所のアンカー場所に自社と競合会社Aおよび競合会社Bを配置し、比較したいお互いの特許群のアンカー図とスケルトン図を描かせて比較参照する機能がある。また、創知のχLUSで自社と競合ベンチマーク2社の選ばれた特許群をレーダ図で色づけ比較する方法がある。その他にもNRIサイバーパテントのTrueTellerに3社の選ばれた特許群だけでヒートマップ図を書かせる比較法などがある。その他にも色々あるが、まだ、各社の一般的な特許検索システムにおいて前述のFTO機能は標準機能にはなっていないと判断する。

IPI-ConfEX2008におけるHunbert Kiehl氏とRenaud Garat氏のプレゼンテーション「Questelの事



例研究」¹⁵⁾は、FTO機能のイメージ作りに参考になった。各部門担当者のコメントも重要視する調査のやり方を示唆している。

6.7 普及容易性

次に欲しいのは検索技術を周囲に広めることを助ける機能である。具体的には3つ考えている。

(1) 検索式のハンドリング容易性

ベテランサーチャーの作成した検索式が保存でき電子メールで簡単に研究者に送れて、全く同じ検索式で研究者が再現できることである。その後で更に研究者が自分で検索式を変更できることが重要である。

(2) 履歴保存性

検索式の作成には集中力と慎重性が求められ検索式の確認も含めて1時間ほどかかり、スパイラル的に検索の質レベルが改善される。人間の弱点は集中力の持続が出来ないのと、記憶力で細かい所が明確に記憶できない点である。そこで支援に威力を発揮するのが検索式と結果の履歴保存の機能である。

(3) Free Viewerという機能

筆者が現在個人的に良く利用している検索システムは日立情報システムズのShareresearchである。そのリンク機能付き代表図付きEXCELファイルダウンロード機能を多用している。作成されたEXCELにはシステムアクセス認証とは関係なく抄録と明細書全文に容易にアクセスでき、ベテランサーチャーの結果を容易に研究者におくることができる。また、創知のχLUSのレーダ図とそのデータの一式はHTMLとして保存され、χLUSの会員にはアクセス認証には関係なくデータ全てがURLを知らせるだけで配信できる機能があり、普及に際して大変便利な機能である。

その他にも欲しい機能はあるが、筆者の判断で主な機能を7つに絞り列挙した。

一方、サーチャーの素養としては6つの役割を務めら

れる能力¹⁶⁾が必要と考える。6つの素養は①技術好奇心・技術理解力、②検索技術力、③知財専門力、④ITリテラシー能力、⑤分析・解析力、⑥コミュニケーション・論理的思考力などである。Japioの創立20周年記念誌とJapio 2007 YEARBOOKと今年のこの記事において、観点は少しずつ異なるが「特許情報検索システムに期待される機能（必要な欲しい機能）」をそれぞれ毎年7つ列挙した。併せて参照して頂けたら幸甚である。

7 おわりに

特許情報の検索システムはどこまで改善が必要か。筆者は永遠のテーマと考える。検索システムを利用者が使えば使うほど、そのシステムの仕組み構造と機能を理解できるようになり、「こんな機能が有ったら良いなあ〜」と感覚が敏感になり、気づきに成長するからである。だから、ヘビーユーザが「素晴らしい気づきと改善案」を持っている。往々にして言わないだけである。一方、プロバイダー側はシステム設計を「こうなるはず」の建て前の企画の視線で考えるために利用者の経験から来る改善案を容易にはイメージアップ連想できない。だから、ユーザの意見に耳を傾けるしかないのである。ユーザと検索システム設計者との協働作業が今後とも将来にわたり求められている「改善の王道であり常道になる」と考える。

筆者は、日本特許庁の最適化計画ロードマップ¹⁷⁾に掲載されている平成22年までが官民あげて特許検索システムをどの様に改善するか考える「ひとつの節目の機会」と考える。その一里塚になる平成20年から21年は特に重要な大事な時期にある。

参考文献

- 1) <http://www.ipi-confex.com/>
- 2) <http://www.piug.org/2008/an08meet.php>
- 3) <http://www.epo.org/about-us/events/emw2008.html>
- 4) <http://www.epo.org/about-us/events/pi-conference2008.html>
- 5) <http://www.sciencedirect.com/science/journal/01722190>
- 6) http://www.infosta.or.jp/pat_d/index.htm
- 7) Conference report; IPI-Confex congerence and exposition, Seville, Spain, March 2008. WPI Vol30 (2008) 253-254.
- 8) The future of European patent information. Curt Edfjall. WPI Vol30 (2008) 135-138.
- 9) Half a century of the Patent Documentation Group (PDG) 1957-2007. Minoo Philipp, Bob Appleton. WPI Vol29 (2007) 148-153.
- 10) 特許情報検索技術の向上にむけて～特許検索競技大会のFB～. 桐山 勉、関西特許情報センター振興会、機関誌、No.21, 19-40.
- 11) INIDコードの日本語解説；
<http://ipbase.cool.ne.jp/patentcity/inid.htm>
- 12) I've never seen in the market. Paul Schwander; WPI. Vol22 (2000) 147-165
- 13) 概念の三角形；情報システム・データベース構築の基礎理論、フーグマン著、情報科学技術協会発行、P44.
- 14) INFOPRO2008にて発表予定。A13発表。
- 15) IPI-ConFEX2008の発表。Case Studt: Increasing Productivity & Communication Using Questel's IP Workflow Management Solution. Hibert Kiehl, Renaud Garat.
- 16) サーチャーの6つの役割と顔、桐山 勉；Japio 2007 YEARBOOK (特許情報研究所、設立記念版)、(2007) p108
- 17) 最適化計画のロードマップ、行程表。
http://www.jpo.go.jp/cgi/link.cgi?url=/torikumi/system/system_optimize_re.htm p32