

テキストマイニングによる特許・論文の分析 —特許情報の多様な活用に向けて—

企業において事業の選択と集中、および新事業の創出が大きな課題となるなか、事業と密接な関係にある特許情報の分析に、近年大きな進歩がみられる。本稿では、NRIサイバーパテントが提供する「TRUE TELLERパテントポートフォリオ」の紹介を通じて、テキストマイニング技術を使った特許情報分析の有効性や活用の可能性について考察する。

効率化が課題の特許調査

企業の研究開発部門や知財部門が、特許情報から技術トレンドや自社・他社の技術領域の違いを把握しようとするとき、これまでは各国特許庁が特許に付与する特許分類コードなどを統計処理する方法がとられてきた。こうした「特許マップ」と呼ばれる分析は、専

用ツールなどにより簡単に行える利点がある反面、分析結果が特許の専門家以外にはわかりにくく、特許分類コードが自社の製品や技術の切り分けと一致しないという問題もある。

より詳細に特許情報を分析する場合には、数十ページにも及ぶ特許公報を1件ずつ読み込む「めくり」と呼ばれる作業が必要である。分析対象が数千～数万件になると、意思決定

図1 コレスポンデンスマップの例(北陸三県における企業と大学関係者の出願テーママップ)

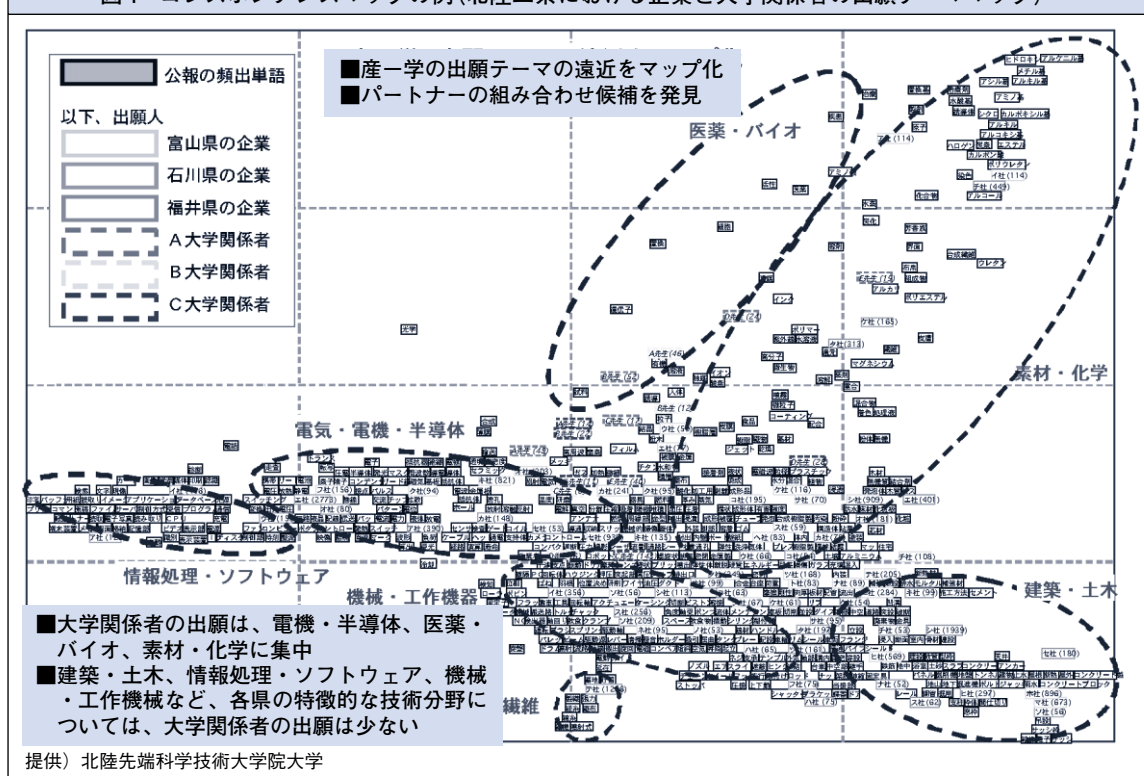
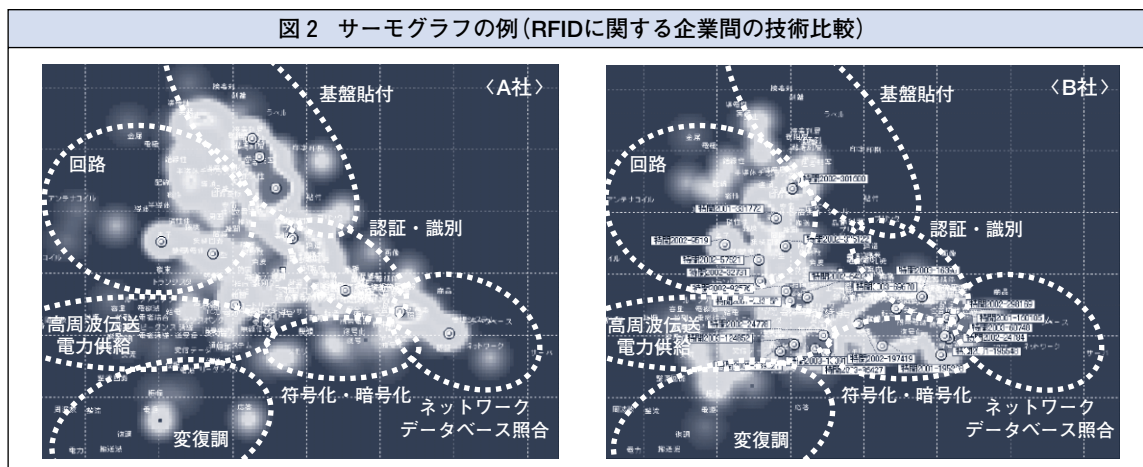




図2 サーモグラフの例(RFIDに関する企業間の技術比較)



の材料を得るまでにあまりに時間がかかり、
分析を断念せざるを得ないこともある。

テキストマイニングによる特許情報分析

こうした状況にあって、大量の特許公報の
記載内容にまで踏み込んだ分析ができる、テ
キストマイニング技術を用いた分析手法が注
目されている。NRIサイバーパテントでも、
野村総合研究所（NRI）のテキストマイニン
グツール「TRUE TELLER」を特許分析に
特化させた「TRUE TELLERパテントポー
トフォリオ」（以下、「TTL-PP」）の販売を
2005年から開始し、多くの企業・大学・研究
機関への販売および受託調査を行ってきた。

「TTL-PP」は以下の2つの分析手法を備え
ている。

①マッピング分析

「TTL-PP」のマッピング分析は、特許公報
群に現れる言葉に基づいた独自のマッピング
手法により、分析対象とする技術分野におけ

る企業の位置や技術の広がりや、短時間で簡
単に可視化するものである。

図1のコレスポネンシスマップはマッピン
グ分析の1つで、特許を出願した各企業・大
学関係者（出願人）と、企業が使用する用語
の関係を1枚のマップに表現することで企業
の位置付けを可視化する。この例では、北陸
三県における企業間連携や産学連携の可能性
を把握することなどが容易になる。

図2はサーモグラフの例である。はじめに、
母集団の文中に頻出する技術用語について、
一緒に使われることの多い単語同士が近くな
るように自動配置し、単語の分布による技術
領域（「回路」「高周波伝送・電力供給」など）
を形成する。これを共通の下敷きとして、企
業別の特許出願の分布密度の広がりや変遷を
サーモグラフ（温度表示）として可視化する
ことで、自社の相対的な強み・弱み、パート
ナー企業との技術的な補完関係を把握できる。
図の例では、A社は「基盤貼付」に集中して

いること、B社は「符号化・暗号化」「ネットワーク・データベース照合」を中心に幅広い技術へ取り組んでいることがわかる。

②グルーピング分析

文献中に出現した用語を、頻度順のリストから選択・分類することにより、文献群を意味的に仕分けし、出願傾向を定量的に分析するものである。図3は、光触媒技術に関する数千件の特許公報について意味的に仕分けして分析したものである。この例では、大学・研究機関では「先端技術」に関する最近の特許が多いこと、企業では「道路・交通」など「用途」を特定した特許に特徴があることがわかる。

従来、企業においては人手でこうした仕分けを行ってきたが、その場合、仕分けの「箱」（「水浄化」「抗菌・防黴」など）をあらかじめ決める必要があり、かつ、手戻りも頻繁に発生した。テキストマイニングを利用すれば、頻度順の単語リストを基に「箱」を設定しな

がら、同時に仕分けを行うことができる。

研究者・組織のつながりを可視化

NRIサイバーパテントでは、「TTL-PP」と組み合わせて、研究者や組織間のつながりを可視化することのできるネットワーク分析ツールを提供している。

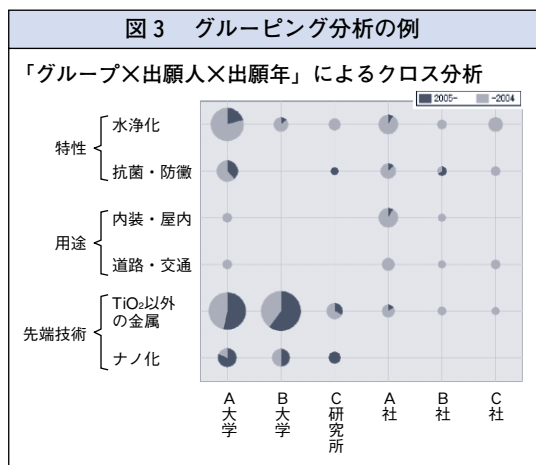
図4は、ある研究機関の特許および論文データに基づいて、それぞれ共同出願、共著となっている研究者・組織のつながりをネットワーク図として表したものである。接続線の上に重ねて表示されたキーワードから、それらの技術的つながりが把握できる。また、時系列でつながりの変遷を可視化することも容易である。

広がるテキストマイニング分析の用途

テキストマイニング技術を用いた特許分析は、企業の技術的な強み・弱みの分析にとどまらず、さまざまな目的・場面で活用が可能である。以下にその例を紹介する。

①自社の要素技術（素材・部品）の用途探索

多くの企業が新事業の創出に苦慮するなかで、既存の要素技術について、特許情報をヒントに新たな事業機会を模索するものである。たとえば、自社が保有する素材の特長（例：耐熱性）などをキーワードとして、まず技術分野の異なる特許まで幅広く抽出する。ここから、テキストマイニング技術を用いて技術の用途に関する言葉を抽出することで、想定



していなかった分野への適用や、新たな技術提携・提供先のヒントを得ることができる。

②取引先のキーパーソンの把握

素材メーカーや部品メーカーがセットメーカーに自社製品を営業する場合などに、相手企業の技術的なキーパーソンに当たれないというケースは少なくない。「TTL-PP」とネットワーク分析を活用すれば、相手企業の特許情報からキーパーソンを見付けたり、競合他社と相手企業とのつながりまで把握したりすることもできる。

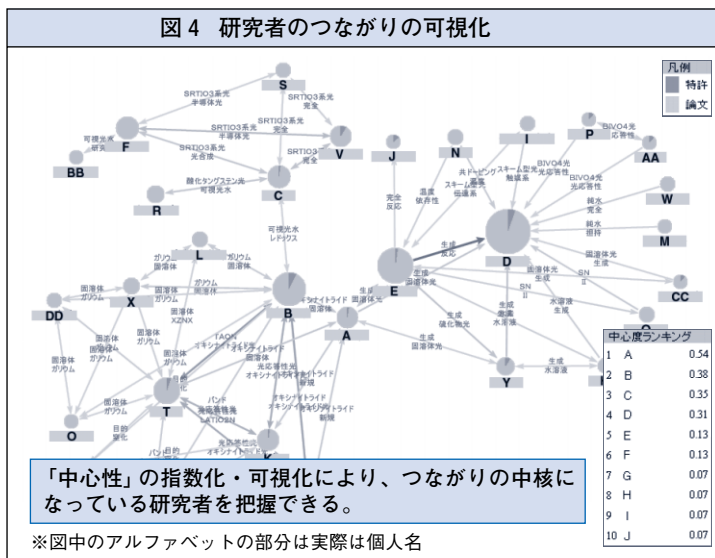
③特許・論文の統合分析

特許情報に加えて論文情報を分析対象に加えることで、特許出願が少ない大学や研究機関での最先端の基礎研究の動向などを幅広く把握することが可能となる。特に、医薬・化学・素材分野の基礎研究や、基盤技術の分野では、論文情報の分析は利用価値が高い。また、特許の場合、出願されてから情報が公開されるまでに制度上、1年半のタイムラグがあるため、論文情報はこれを補完する上でも有効である。

④産学連携の候補先の発掘

産学連携は、学会などを通じた個人的なつながりや、技術雑誌などでたまたま見付けた記事を契機としたつながりによることが多い。企業における基礎研究の継続的な取り組みが

図4 研究者のつながりの可視化



「中心性」の指数化・可視化により、つながりの中核になっている研究者を把握できる。

難しくなるなかで、特許・論文情報から幅広くキーパーソンを探索できることは、連携のためにコンタクト先を見付ける上で有効である。大学や研究機関においても、研究内容の事業化は課題となっており、パートナー候補を広く探索できることは有益である。

⑤技術政策・産業政策の効果測定

特許・論文情報の解析は、公的研究機関による研究開発プロジェクトがどれだけの知的財産を生み出し、その後の民間企業などによる技術開発にどう波及効果を与えているかを分析するためにも有効である。また、特許・論文の件数比較から、日本が優位な技術領域を見極め、今後の技術政策や産業政策に活かそうといった議論も行われている。

以上のようにテキストマイニングを用いた特許分析の活用は、今後ますます広がっていくものと思われる。