

テキストマイニング手法の有用性の検討

知的財産情報検索委員会
第 1 小委員会*

抄 録 特許の分析対象特許件数が増大し、効率化が求められる中、特許公報中の文章をテキストマイニングにより分析する技術が登場し、様々な分析ツールに取り入れられてきた。テキストマイニング分析は、言葉で技術を把握できる点、短時間に大量のデータを可視化できる点で、書誌分析やマニュアル分析等の従来の分析手法と比べて優れている一方、従来の手法に比べて結果の妥当性の把握が難しく、広く使いこなせていないという印象もあった。当委員会では、テキストマイニングツールの「活用」に着目し、活用シーンとツールの特性を理解することを通じ、テキストマイニングが有用な分析手法となりうることを確認した。

目 次

1. はじめに
2. 目的・課題
3. テキストマイニングの仕組み
4. 活用シーンを想定したテキストマイニング分析
 4. 1 ベンダーへのヒアリング
 4. 2 分析事例紹介
 4. 3 マップの種類と好適シーン
 4. 4 特許分類による分析との比較
5. まとめ
6. おわりに

1. はじめに

「テキストマイニング」とは、文書情報（自然言語で書かれたテキスト情報）から有益な知識を発見・抽出するための技術である。情報の電子化に伴い大量に蓄積されるデータから新たな知見を発掘（マイニング）する分析方法であり、データマイニング手法の一種である。従来データマイニングが扱うのは数値データであったが、電子化されたテキスト情報もマイニング対象とするのがテキストマイニングであり、文書を個別に調べてもわからない、文書群に内在

する知識を発見することを目的としている。

従来、特許の技術的な分析は、特許公報に付与されたIPC(国際特許分類)、FI(File Index)、Fターム (File Forming Term) をはじめとする特許分類、またはマニュアル付与した分類に基づき行われてきた。2000年を過ぎて、新たな分析手法として特許公報中の文章をテキストマイニング手法により分析する技術が登場し、従来よりも短時間で大量のデータを分析ができる手法として新たな分析手法の選択肢の一つとなった。

しかし、調査担当者にとって、テキストマイニング分析は、従来の手法に比べて結果の妥当性を把握したり効果を評価することが難しく、広く一般ユーザーを対象とする本格的な「活用」に関してはまだ課題が残されていた。

当ワーキングの活動では、テキストマイニングツールの「活用」に着目し、活用シーンとツールの特性を理解することを通して、テキストマイニングの有用性の検討を行うこととした。

* 2010年度 The First Subcommittee, Intellectual Property Information Search Committee

2. 目的・課題

本研究の目的・課題を明らかにするため、テキストマイニングの使用経験を有するメンバーからテキストマイニングツールに対する感想および意見をヒアリングした。

ユーザーがテキストマイニングツールに抱いていた期待とは、具体的には、「膨大なデータを短時間で技術俯瞰したい」、「対応分類がない分野／分類精度が低い分野を調査したい」「新しい気づきが欲しい」「視覚に訴えたプレゼンがしたい」等が挙げられた。一方、実際に導入またはトライアルをして使ってみた感想としては、例えば、調査報告などで視覚に訴えたテキストマイニングの特許マップは好評であったものの、「データの解釈が難しく説明に自信が持てない」、データの理解のために「読まないつもりでいたが結局1件ずつ読むことになり時間の短縮が図れなかった」、また「テキストマイニングツール導入効果の説明が難しい」という意見が挙げられた。

ヒアリングの結果から、テキストマイニングツールに対するユーザーの期待と現実に使ってみた感想の間にギャップがあることが明らかとなった。

これらのギャップは、ユーザーの理想とする方向にツールの機能や性能が向上および改善することにより埋まり得るものかもしれない。しかしながら、かかるギャップはユーザー自身のテキストマイニングツールに対する考え方・使い方にも起因するのではないかと考えられた。つまり、ユーザーの使い方が変わることで、従来のテキストマイニングツールをより適切かつ有効に活用できるのではないかと考えた。

そこで、現在利用可能なツールの特性と、分析の活用シーンを理解することで、ユーザーの視点からテキストマイニングツールの有用性を検討することとした。

3. テキストマイニングの仕組み

上述のように、従来の特許分類を用いた特許分析では、特許文献に付与された特許分類および他の書誌事項を分析項目として、単純な定量分析やクロス集計¹⁾を行うことにより分析することが多い。分析結果の理解においては、分析項目である特許分類記号のシステムと記号が示す意味を理解した上で、分析結果を解釈する必要があった。

一方で、テキストマイニング分析では、分析項目が単語で表されることから、分析項目は視覚的に認識しやすく、特許分類の知識のないユーザーにも親しみやすい。しかし、単純集計やクロス集計などのシンプルな分析手法に比べ、データ入力から可視化に至るまで複雑な分析過程を経ていることが、分析手法に基づき結果(多くはマップ)を解釈する際にユーザーを悩ませる一因となっていた。

まず、テキストマイニングツールの活用への1段階目として、一般的な特許のテキストマイニングツールの言語処理手法を、テキストマイニング手法に関する文献およびツールのベンダーへのヒアリング(後述)を通じて、図1のような単純化した模式図に落とし込んだ。

一般的なテキストマイニングツールにおいて、文書が可視化されるまでのプロセスを図1の①～⑥に示す。

① 特許データベースから取り込まれた請求項や要約等のテキスト情報から、形態素解析²⁾を経て、単語が切り出される。

② 切り出された単語から、技術の特徴を表す重要な単語である「キーワード」が抽出される。代表的なアルゴリズムとしてTF-IDF法³⁾が挙げられる。キーワードは、完全自動で抽出される場合もあれば、自動で抽出された候補群からユーザーが抽出する場合もある。なお、ノイズ語処理⁴⁾、同義語処理⁵⁾はキーワードの抽

出の前に行われるが、実際はキーワードの抽出結果からフィードバックする形で反復的に行われる。

③ ②で抽出されたキーワードの出現頻度や共起性⁶⁾に基づき、文書間の意味的距離（類似度）を尺度化する。代表的な尺度化方法としてベクトル空間法⁷⁾が挙げられる。

④ ③で尺度化した結果に基づき、類似文書ごとにグループ化する（クラスタリング）。代表的な尺度化方法として最短距離法、最長距離法、群平均法⁸⁾が挙げられる。

⑤ ④の処理で得られたデータが多次元情報である場合、2次元のマップに表示できるように情報変換を行う。

⑥ ⑤の処理で得られたデータをマップとして可視化する。

なお、一般的な特許分析ツールには、従来型の特許分類や出願人名等の書誌的事項の統計分析機能（以下、「書誌分析」と表す）が備わっていることや、テキストマイニングと併用できることが多いので、かかるプロセスも左側に併

記している。

かかる模式図は、言語処理技術・数学分野に馴染みの薄いユーザー、特に当該分野にアレルギーのあるようなユーザーでも、ツールの利用に必要最低限のテキストマイニング手法の特徴を理解することを容易にする。そして、模式図を通じて、一般的なテキストマイニングの言語処理を単純化して理解するだけで、テキストマイニングツールに対する意識を変えることができる。

例えば、テキストマイニングの言語処理を一般化の形で理解しておけば、ブラックボックスとして解釈していた時よりも遥かに適切な解釈ができる。また、形態素解析で文章から切り出した無数の単語の中から、技術の特徴を表す重要な単語として抽出された「キーワード」に基づく分析がなされることから、テキストマイニングは全ての語を考慮した意味分析ではないことが容易に理解できる。すなわち、テキストマイニングツールに目視と同等の文書識別力を求めるのは言語処理メカニズムから考えても過度な期待であることが明らかになった。

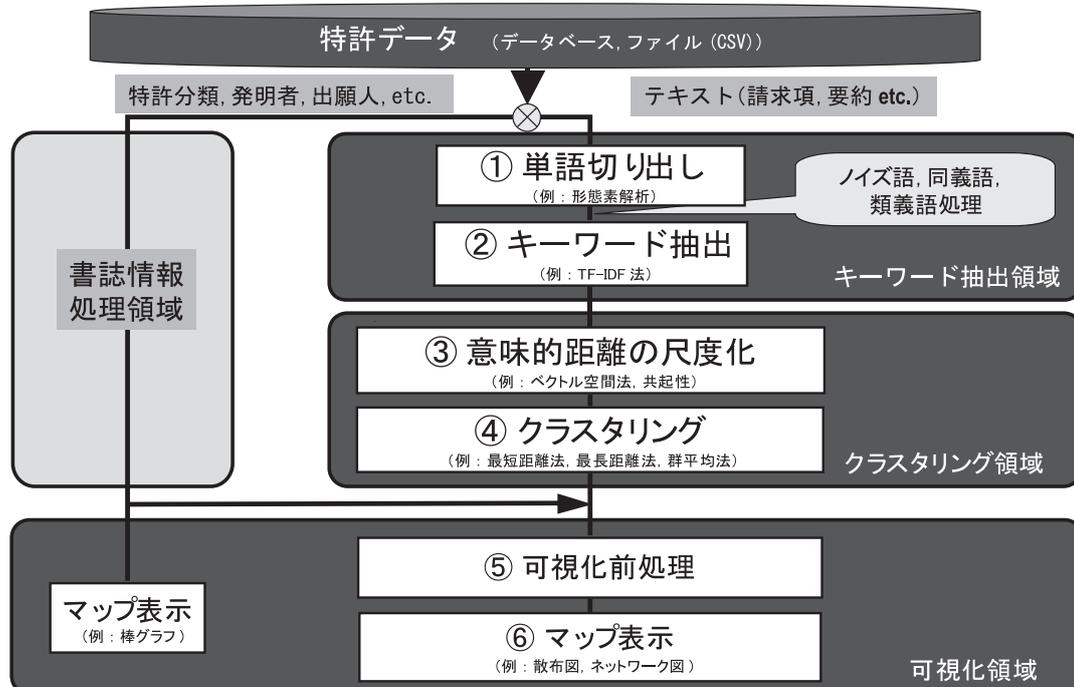


図1 一般的な特許情報のテキストマイニングのプロセス

4. 活用シーンを想定したテキストマイニング分析

活用方法を見出すためのアプローチとして、実務で起こりうる活用シーンを想定した場合のテキストマイニングマップの使い方から、マップの特徴、およびかかる特徴に基づく活用への手がかりを探ることとした。

具体的には、平成15年および17年の特許流通支援チャート「LED照明」⁹⁾ で用いられた検索式を参考にして作成した集合¹⁰⁾ に対して、以下の5つの活用シーンを想定した分析を行うこととした。

- ① 事業戦略立案：アライアンス先検討，新規事業探索
- ② 開発戦略立案：課題・解決手段の相関
- ③ ニーズ調査：素材や要素技術の用途探索
- ④ 先行技術調査：類似技術の把握
- ⑤ 技術動向調査：業界間，業界内のプレイヤーの関連性

テキストマイニングツールとしては、「日本語を分析対象とすること」，「5,000件以上（目視が困難な程度の数を想定）の特許を分析可能であること」を条件に選定し，以下の5つのツールを検討対象とした（アルファベット順）。

- ・ ATMS/Analyzer（富士通）
- ・ Biz Cruncher（パテントリザルト）
- ・ REXION PRO（山武）
- ・ TRUE TELLERパテントポートフォリオ（NRIサイバーパテント）

- ・ χ LUS（創知）

研究に用いたテキストマイニングツールの概要を表1に示す。表中，タイプAは，複数のテキストマイニング分析機能と，従来の書誌分析ツールの機能を並存させたタイプであり，タイプBは，単一のテキストマイニング分析機能を有し，そのテキストマイニング分析機能を補完する形で書誌分析機能が存在するタイプである。

4. 1 ベンダーへのヒアリング

まずは，テキストマイニングツールのベンダーのご協力のもと，ベンダーの視点からの前述の各シーンにおけるテキストマイニングマップの使い方や特徴の傾向から活用への手がかりを探った。

一例として，図2にTRUE TELLERパテントポートフォリオ（NRIサイバーパテント）による分析例を挙げる。

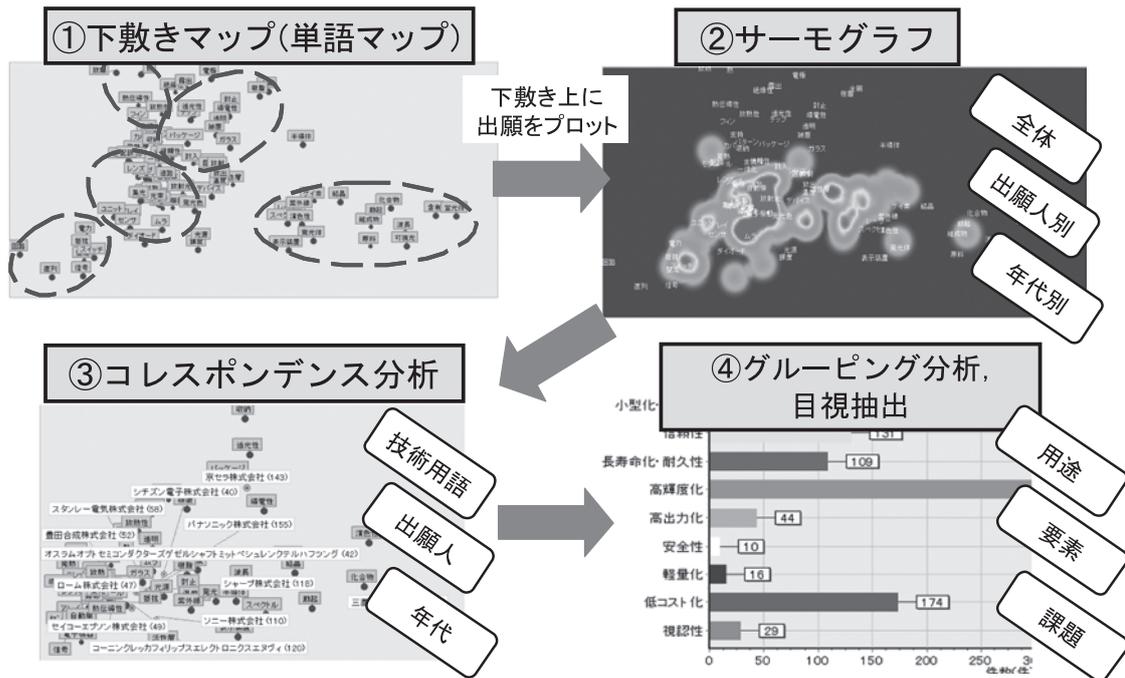
マップの種類，言語処理アルゴリズム，マッピングされたテキスト，併用されたマップ等の情報を，各ツールに対して図2のように整理したものを比較し，活用シーンとテキストマイニングマップ間の相関関係やパターンを見出し，活用方法へのヒントを得ようと試みた。

まず，活用方法の全体的な特徴として，以下の2点が挙げられる。

- ・ 複数のテキストマイニングマップを搭載したAタイプのツールではマップの使い分けに特徴が見られた。また，従来の書誌分析と併用がなされ，特許分類とテキストマイニングによる分

表1 研究に用いたテキストマイニングツールの概要

ツール名	タイプ	ツールの保有するテキストマイニングマップ例
ATMS/Analyzer	A	アンカーマップ，スケルトンマップ
Biz Cruncher	A	課題・解決マトリクス，製品分類，構造図
REXION PRO	B	樹形図
TRUE TELLER	A	サーモグラフ，コレスポネンス分析
χ LUS	B	レーダー図



(書誌分析は概況として最初に提示され、シーン別の説明はテキストマイニング中心で行われた)

シーン	分析方法 ○：テキストマイニングを用いた分析，△：テキストマイニングに関連した分析，×：テキストマイニングとは異なる分析
売却先	○ サーモグラフによる企業間比較（合算と差分） ○ コレスポネンデンス分析による企業間比較
課題・手段	○ グルーピング分析（課題×解決手段）によるマトリクスマップ（係り受け使用）
用途	△ 対象テキストに利用分野と従来技術も加えて、自動グルーピングによる用途抽出（係り受け使用）
類似特許	○ サーモグラフ上での対象案件のプロットと近傍案件抽出 ○ 係り受け等による課題 x 解決手段のクラスタリング
全体動向	○ 出願人ごとのサーモグラフから競合先を選定 × ネットワーク分析から潜在的な競合先を選定 (⇒その後、対象企業の動向を様々な観点で調査（省略）)

図2 TRUE TELLERパテントポートフォリオ（NRIサイバーパテント）による分析デモの概略

析は、それぞれ異なる項目・観点で行われた後に、両者の結果を踏まえて最終的な結論を導き出していた。例えば、全体の動向を把握するに際し、特許分類では分類の区分に沿って技術の分布がマップ化され、テキストマイニングでは特許やキーワードのマップ上の分布状況から技術の分布を把握し、両者の情報に基づき全体分布が解釈される。

・Bタイプのツールは基本的に、Aタイプのツ

ールでも場合により、一つのテキストマイニングマップを、シーンに応じて、分析対象テキストや着眼点を変化させて対応させていた。前者の分析対象テキストの変化の例としては、課題の抽出では「課題」のテキストを分析対象としたり、用途では「技術分野」を分析対象とすることが挙げられ、後者の着眼点の変化としては、要約を分析対象とする場合に、キーワード候補から「課題」「構成」「ある技術に関連する用語」

をシーンに応じて選択して、課題や解決手段、技術の傾向を把握することが挙げられる。

一方、マップの種類と活用シーン、言語処理アルゴリズムと活用シーンの間においては、着眼点をはじめとする主観的要素や、マップの併用の影響もあり、単純かつ有意な相関を見出すことができなかった。

以上より、テキストマイニング分析は書誌分析に代わる分析手法というよりも、従来から行われている書誌分析と併用することで、従来とは異なる多様な観点の分析ができると考えられる。

4. 2 分析実例紹介

次に、テキストマイニング分析を行い、前述のように、テキストマイニング分析がもたらす従来の書誌事項のみの分析とは異なる有用な情報に着目して検討を進めた。分析結果の代表例を図3～6に示す。

図3：スケルトンマップ(ATMS/Analyzer(富士通))

全体集合ではキーワードが分散し特徴となる語が埋もれてしまう傾向が見られたが、A社の集合に絞り込むことで、A社特有の課題が明確となった。特に、マップ中で示すように、「光取り出し効率」や「演色性」など、「LED照明特有の課題」、「Fタームにはない課題」を、ラベルから容易に見つけることができた。なお、図3のマップはマニュアルで同義語処理を行った後のものである。

図4：課題・解決マトリクス(Biz Cruncher(パテントリザルト))

マニュアルでノイズ語処理、同義語処理を行い、マップを作成した。全体集合では特徴的な課題を見つけることが難しかったが、集合を「5F041 AA43 信頼性の向上」に絞り込むことで、信頼性に関連する細かいキーワードを見出すことができた。

図5：樹形図(REXION PRO(山武))

マニュアルでノイズ語処理、同義語処理を行った後、マップを作成した。全体集合では技術内容が分散し、集合を1つの樹形図にマップ化することも、特徴を見出すことも難しかったが、素子メーカー3社の出願からなる集合を樹形図で可視化すると、技術毎に枝(クラスタ)が形成され、各社の出願の特徴、A社B社の出願領域に偏在があることとC社が全領域満遍なく出願していること、を視覚的に把握することができた。

図6：レーダー図(χ LUS(創知))

自動的なマップ化処理により、全体集合が1つのレーダー図上に可視化された(密度の高い領域を切り取り図示した)。会社毎にクラスタをハイライトし比較することで、各社の出願の技術分布を俯瞰的に把握することができた。また、クラスタ毎の類似距離がマップ上の距離と対応することから、座標データに基づき各社の年毎の座標の重心の推移をプロットし、レーダー図と比較することで、「全体的に応用分野に移行している」等、この分野における各社の開発の方向性を把握するための手がかりともなりうる。

以上のようなワーキングによる分析結果から以下のようなテキストマイニングならではの長を挙げることができる。

- ・特許分類にはない観点(新しい観点、深い観点)で分析できる。
- ・分析文書数の規模が大きい中でも個々の変動を見ることができる。
- ・読む前に集合の概要をキーワードにより把握できる。

さらに、上述の有用な機能を発揮させるために「キーワード」の取り扱いが重要であることがメンバーの分析作業からの気づきとして得られた。例えば、キーワードの選択や解釈、シソーラスの整備状況によって結果が変わるからである。つまり、ユーザーは分析対象技術、キーワードの処理ができ、キーワードが対象技術において果たす役割を理解できる程度の予備知識

A社の課題の掘り下げ

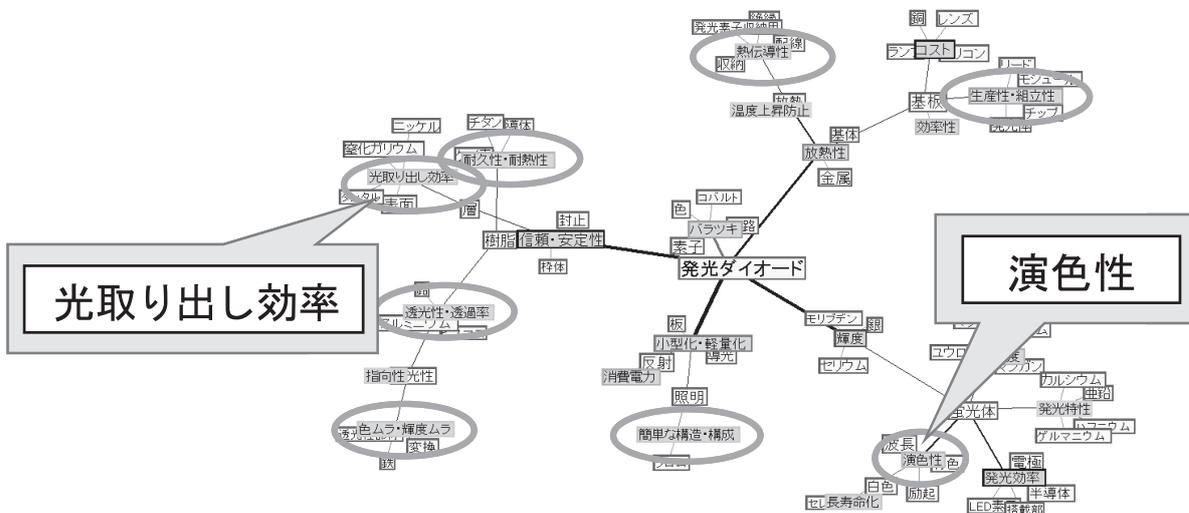


図3 スケルトンマップ (ATMS/Analyzer (富士通))

Fターム内の課題の掘り下げ



図4 課題・解決マトリクス (Biz Cruncher (パテントリザルト))

技術分布の視覚的な把握

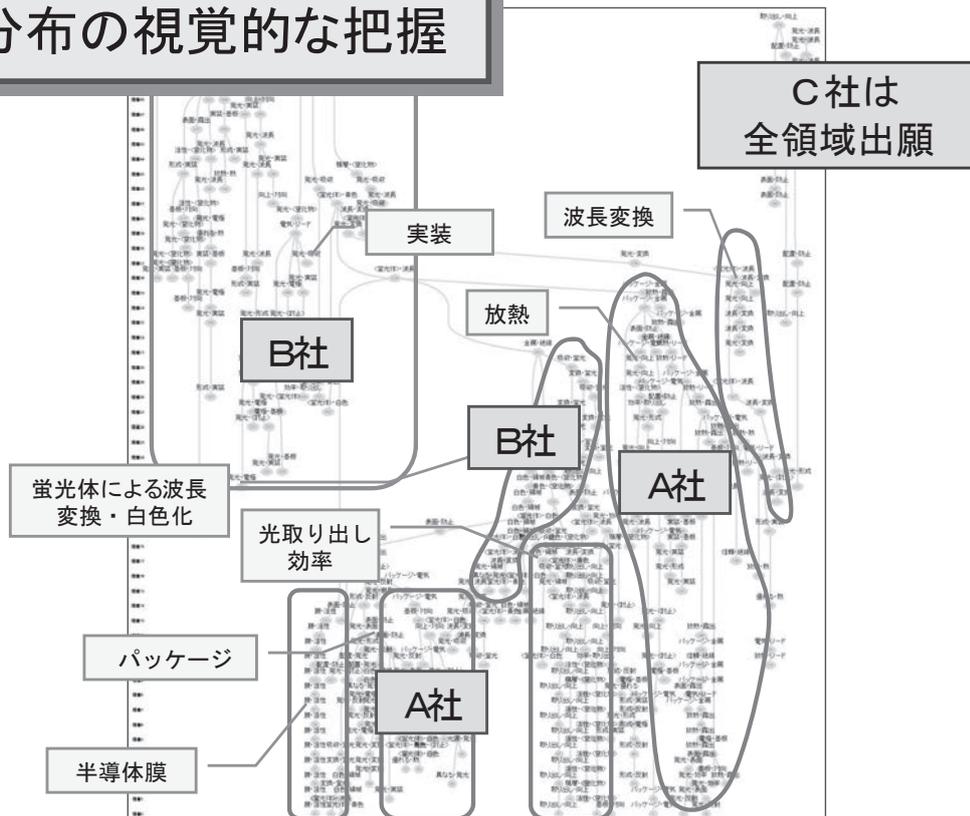
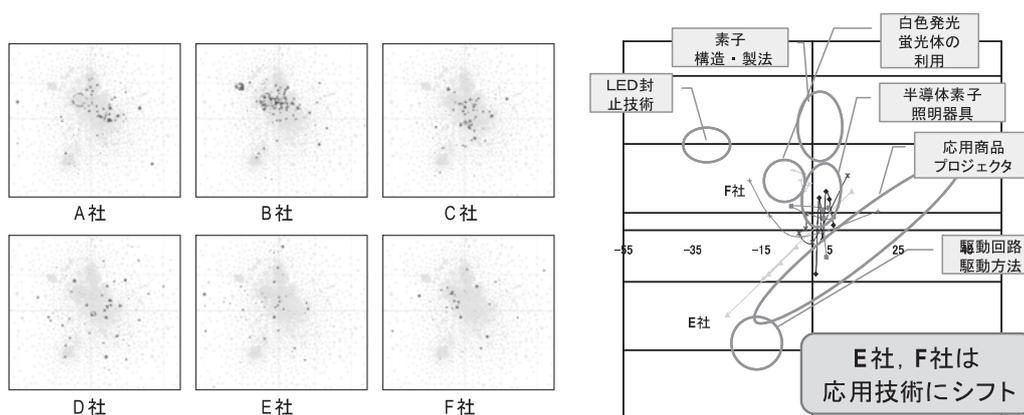


図5 樹形図 (REXION PRO (山武))

技術の推移を俯瞰的に把握

照明用LED素子メーカー6社 2005年~2009年(公開年)出願分布



レーダー図

重心解析

図6 レーダー図 (X LUS (創知))

を有することが、妥当な結果を得るための基本的な条件であると考ええる。

4. 3 マップの種類と好適シーン

テキストマイニングマップにより、一度にマップに可視化できる文書量またはキーワード数が異なり、マップの表す事象も異なる（例えば、文書の類似性やキーワード相関性）。一度にマップに可視化できる文章量またはキーワード数が比較的多いマップ、例えば、レーダーチャート（χLUS）やサーモグラフ（TRUE TELLERパテントポートフォリオ）は、集合全体を把握することのできる概要把握向きのマップであり、一方、数個～数十個のキーワードやキーワードに代表される文書の相関関係を示すマップは、ある程度のまとまりを有する集合を深く掘り下げて分析できる詳細把握向きのマップといえる。一般的な言い方をすれば、概要把握はマクロ分析、詳細把握はマイクロ分析としても言い換えることができるだろう。メンバーが実際にテキストマイニングツールを使って実感したことの一つはこのマクロ・マイクロの違いである。特

に今回はLED照明とはいえ、半導体レベルから照明器具レベルまで広範な技術を包含した集合を用いたため、特に顕著にこの違いと、それに起因する情報の質の違いが実感された。そして、マクロ・マイクロの違いと、活用シーンの相関に傾向があることはベンダーのヒアリング、メンバーの実際の分析を通じて把握された。これを図7に示す。

4. 4 特許分類による分析との比較

前述のとおり、テキストマイニング分析は、従来の書誌分析と並存する手法であると述べた。特許分類を用いた分析においても、マクロ分析—マイクロ分析という考え方が一般的になされており、テキストマイニング分析と共通する。そこで、本研究を通じて明らかになったテキストマイニング分析と特許分類による分析の特徴を表2に比較表として示す。

表2に表す特徴に加え、特許分類がある程度成熟した技術または一定数の出願のある技術に対して作成されることを踏まえると、テキストマイニング分析は、特に、特許分類の枠組のな

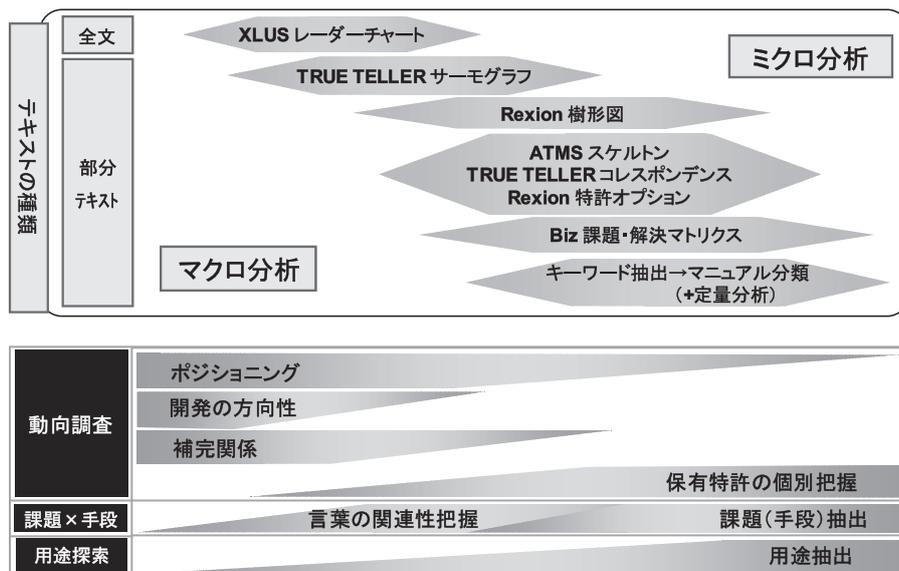


図7 特許マップと活用シーンの相関

表2 テキストマイニング分析と特許分類による分析の比較

	テキストマイニング分析	特許分類による分析
文書またはクラスタに付与されるラベル	文章から抽出した「キーワード」	分類記号 (IPC, FI, Fターム)
文書またはクラスタの解釈	1つまたは複数のキーワードに基づき、ユーザーが解釈を行う	分類の定義に従う
利点	ラベルによる直感的な把握が可能 特許分類とは異なる観点、深い観点で分析できる 大きな集合の中の小さな変動も可視化できる	文書またはクラスタの解釈が容易 ラベルが集合により変動しない
欠点	文書またはクラスタの解釈がユーザーに委ねられる ラベルが集合の質により変動する	分析の観点が分類の枠組みに制限される 大分類を用いた分析では集合内の小さな動きが見えないことがある
分析方法	全体把握→相関分析→キーワード抽出・定量	FI・IPC・テーマコード→Fターム分析 (→マニュアル分類分析)

い新しい技術や特殊な技術、あるいは分類の区分を超えるような技術の分析の場で有効であるう。

5. まとめ

テキストマイニング分析が文章から抽出した「キーワード」を用いた分析であり、従来の分析手法とは異なる観点で分析できる手法であるが、従来の分析手法に取って代わるものではなく、併用することにより精度の高い分析が期待できる手法であった。

また、テキストマイニングはそのマップに可視化できる情報量によりマクロ分析向きの分析と、ミクロ分析向きの分析とに分けられ、活用シーンとそこで必要な分析の種類に応じて、ユーザーはテキストマイニング機能を選択する必要がある。

なお、テキストマイニングツールを使えば、特許公報の目視は不要になるという期待を人によっては抱きがちであるがその言語処理プロセスから、その期待が大きな誤解であることは明らかである。さらに、未知の分野でもテキストマイニングを使えば容易に分析できるという期

待も誤解である。可視化に至るまでのプロセス（例えば、同義語・ノイズ語処理）、ラベルの解釈、および可視化された情報の理解には、ユーザーの「解釈」が特許分類を用いた分析と同じく重要な役割を果たすことから、対象技術への知識を有することは分析の前提条件といえるだろう。

加えて、従来の分析手法と同じように、調査目的を明確にすること、集合の質やサイズを分析に適したものにすること、考察にあたり仮説を立て検証しながら進めることが必要であると実感した。

6. おわりに

今回、テキストマイニングをもっと活用するためにはテキストマイニングとどのように向き合えばよいのか、というユーザーの視点から検討を行った。しかしながら、より良い分析・解析のためには、ユーザー自身の意識、行動に加えて、ツールの進化も必要である。

最後に、ツールに対する要望も述べたい。まずは、キーワードの抽出でユーザーの関与が重要となるが、より効率よく分析を進めるため、

ノイズ語および同義語辞書、特に特許文書向け辞書、技術分野別辞書の開発と搭載が望まれる。また、どのテキストマイニングマップにも最適な活用シーンが存在するが、実際の企業での幅広い利用シーンを考えると、複数のテキストマイニングマップが搭載されることを希望する。さらに、日本特許のFI、Fタームの分類項目は多く、付与精度も比較的優れていることから、特許分類との併用も有効であると考えられる。また、調査目的に応じたテキスト項目選択が自由かつ容易に行えると好ましい。

本稿で言及されたテキストマイニングツールの情報は2011年3月時点での情報に基づく。

NRIサイバーパテント株式会社、株式会社創知、株式会社パテント・リザルト、富士通株式会社、株式会社山武(敬称略)の各社には、ヒアリングおよびツールの提供にご協力いただいた。

尚、本稿は2010年度知的財産情報検索委員会第1小委員会第2WGメンバーである大石明美(帝人知的財産センター、小委員長補佐)、小松拓也(古河電気工業)、高井史比古(セコム)、谷口誠一(セイコーエプソン)、中村康行(リコー)、山崎勇二(三菱化学)、渡邊芳代子(三井・デュポンフロロケミカル)が担当した。

注 記

- 1) 与えられたデータのうち、2つないし3つ程度の項目に着目してデータの分析や集計を行うこと。1つ(ないし2つ)の項目を縦軸に、もう1つの項目を横軸において表を作成して集計を行う。
- 2) 言語で意味を持つ最小単位に分割する。
- 3) 文書中での単語の頻度TF(Term Frequency)と、単語が出現する文書数の逆数IDF(Inverted Document Frequency)により計算される。「文書中に多く出現する単語」の重みを大きくし、「多くの文書に出現する単語」の重みを小さくする重み付けの指標。
- 4) 発明の特徴を表さない単語を分析対象から除外すること。

- 5) 同じ意味を表す別の単語を1つの単語群としてグルーピングすること。
- 6) ある単語がある文章中に出了とき、同時にその文章中に別の単語が出現すること。
- 7) 各文書を、キーワードの重みを要素として持つ多次元ベクトルとして表現し、ベクトルの内積により文書間の関連度・類似度を計算する方法。
- 8) いずれも距離が近いものからクラスタ化するか、または逆に遠いものを別クラスタに分ける手法であり、最終的にツリー状の樹形図を描くことのできる階層的クラスタリング手法。クラスタ間距離(類似度)の測定方法により最短距離法、最長距離法、群平均法等に分けられる。
- 9) 独立行政法人工業所有権情報・研修館：特許流通支援チャート「電気19 照明用LED技術」
<http://www.inpit.go.jp/katsuyo/archives/archives00007.html> (参照日2012年1月10日)
- 10) 分析には事例に応じて(1)または(2)のいずれかの集合を用いた。

集合(1)

① FI:H01L33/00 階層検索

②「全文」KW検索「照明」

③ ①AND②

⇒8,857件

集合(2)

①:FI:H01L33/00 階層検索

②「要約+請求項」KW検索「照明」

③ ①AND②

⇒3,176件

参考文献

- ・豊田裕貴, テキストマイニングによるドキュメントデータの分析, 情報の科学と技術 Vol.53, pp.22-27 (2003)
- ・渡部勇, テキストマイニングの技術と応用(〈特集〉情報の分析・解析法), 情報の科学と技術 Vol. 53, pp.28-33 (2003)
- ・中居隆, テキストマイニングによる知財ポートフォリオ分析, 情報管理, Vol. 51, pp.194-206 (2008)
- ・村上 英治, 木幡 真望, 産業応用分野に求められるテキストマイニング技術と適用事例-実務を意識した定性情報のクラスタリングと可視化技術, azbilテクニカルレビュー, pp.30-37 (2006)
- ・ローネン・フェルドマン, ジェイムズ・サンガー,

本文の複製、転載、改変、再配布を禁止します。

- テキストマイニングハンドブック, pp.1-23 (2010),
東京電機大学出版局
- ・中村達生, JICST ファイル・特許 DBを用いた動向
分析, 情報管理, Vol. 46 (2003), No. 2 pp.97-106
 - ・那須川哲哉, テキストマイニングを使う技術／作る
技術—基礎技術と適用事例から導く本質と活用法,
(2006) 東京電機大学出版局

- ・特許庁, 知的財産戦略に資する特許情報分析事例集
—特許情報分析事例集, (2010)
- ・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, pp.20-
31 (2009)

(原稿受領日 2012年1月10日)

