

クリアランス調査におけるスクリーニングの研究

情報検索委員会
第1小委員会*

抄 録 特許情報の調査において、「検索」は様々な効率化手法の検討が行われてきた。しかしながら、重要な特許、即ち鑑定等の最終的な方向付けが必要な特許を逃さず集めるための網羅的な検索を行うと、無関係な特許（以下、本稿では「ノイズ」と呼ぶ）が多く含まれることは避けられない。検索結果をうまく仕分け、重要な特許に素早く、少ない労力でたどり着くために、スクリーニング手法が重要であると考えた。本稿では、近年多数提供されてきている特許検索データベースや特許分析ツールを用いたスクリーニング手法を紹介する。

目 次

- はじめに
- 特許調査におけるスクリーニングの定義
- オートスクリーニング
 - 概要
 - 事例紹介
- マニュアルスクリーニング
 - 概要
 - 事例紹介
- 重要な特許の抽出後
- その他ツールの紹介
- 文書管理上の注意事項
- おわりに

1. はじめに

近年、高機能な特許検索データベースが多数提供され、特許情報調査においては、これら特許検索データベース上で各種特許分類やキーワードを用いた機械検索を行うのが一般的となっているが、機械検索のみで適合率（正確性）と再現率（網羅性）を両立させることは難しい。即ち、重要な特許（事業遂行において障害となる可能性のある特許）を見落とさないように広い範囲を検索して再現率を高めると、ノイズが

多数混じり適合率が低下してしまう。特に、クリアランス調査（特許侵害防止調査）においては、侵害リスクを減らすために適合率よりも再現率が重視され、検索結果の特許の集合（以下、本稿では「特許検索集合」と呼ぶ）の数が膨大となる事例が頻繁に生じ、結果的に多大な工数を充てなければならない状況が多々生じる。そのような特許検索集合は多数のノイズを含む可能性が高く、優先度に応じた仕分け（スクリーニング）を行うことで、適切なリソース配分を行うなどクリアランス調査に充てる工数の効率化が期待できる。

ところで、検索式立案手法や各種特許検索データベースの比較検討などの検索手法の研究と比べて、スクリーニングの手法はあまり検討されていないようである。過去のスクリーニング研究¹⁾も、研究者や技術者などを対象としたものであり、知財実務従事者を対象としたものではない。

本稿では、スクリーニングの効率化に注目し、特許検索データベースや特許分析ツールを用いたスクリーニングの具体的な効率化手法を紹介

* 2014年度 The First Subcommittee, Information Search Committee

する。

2. 特許調査におけるスクリーニングの定義

特許情報調査の一連のプロセスを、図1のように定義する。

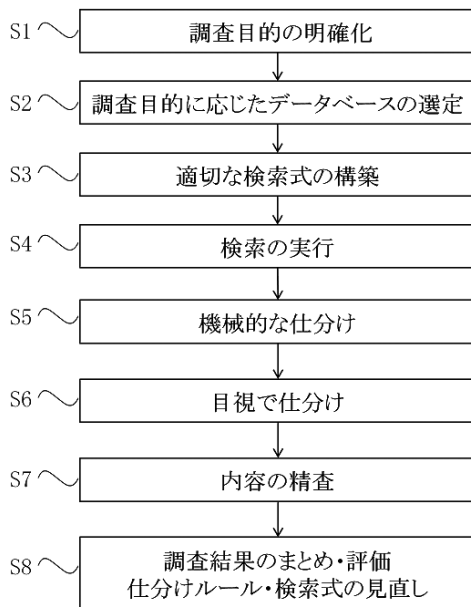


図1 特許調査の概略

本稿で「スクリーニング」とは、検索実行（ステップS4）後に得られた特許検索集合に優先度を付けて部分集合（以下、本稿では「特許束」と呼ぶ）に仕分けることと定義する。即ち、調査すべき特許検索集合を、適合率が高いと見込める部分集合（以下、過去のスクリーニング研究¹⁾に倣い「濃い集合」と呼ぶ）とそうでない部分集合に分けて優先順位付けすることで効率的に読み込むことができる。そこで本稿では、特許検索データベースや特許分析ツールを使用した機械的な仕分け「オートスクリーニング」（ステップS5）と、特許の内容を多少目視で確認して行う仕分け「マニュアルスクリーニング」（ステップS6）を定義し、それぞれ次章以降で説明する。なお、図1に示した特許調査

のプロセスは一例であり、全ての工程を行う必要はなく順番が入れ替わった方が効率的な場合もあるだろう。調査対象の技術分野や特許検索集合の件数に合わせて工夫するとよい。

なお、スクリーニングを効率化するにあたっては、以下に述べることでできていることが前提となるが、本稿ではこれらの工程について言及することを省略する。

A. 調査範囲・調査深度の明確化：知財リスクと調査精度との兼ね合いを考慮する。漏れのない調査を行う必要があるため全ての特許検索集合全体を調べるのが理想だが、現実には全体のうちの一部を調べることも多い。調査報告書には、調査結果と共に「何を調べなかったか」（侵害のおそれはあるが、調査範囲から外したものを）記載することが推奨される場合もある²⁾。

B. 仕分け基準の明確化：調査範囲・調査深度に応じて、スクリーニングにおける仕分け基準や徹底度合いを明確にする。特に、マニュアルスクリーニングとオートスクリーニングを異なる担当者が行うこともあるため、各担当者の役割を明確にする。想定外事象発生時は申し送りや差し戻しを行い調査範囲や仕分け基準を再検討する。

C. 担当者の割り振り：各スクリーニングの工程に担当者を割り振るだけでなく、優先度に応じた各特許束の担当者や人数も調整する。

D. 検索式の精査・改良：スクリーニングを行った結果、検索式の問題点に気付くこともある。そのような場合は調査範囲に見合った検索式に改良する。

3. オートスクリーニング

3.1 概要

特許検索データベースや特許分析ツール上で特許書誌情報などの指標を用いて、ある程度自動的に優先順位を付けることができる。明細書

の内容を読み込まないので、マニュアルスクリーニングの前に特許検索集合を仕分けると効果的である。オートスクリーニングでは、以下に述べる指標を用いることにより優先順位付けすることで、その後の工程で適合率が高い特許束から調査を進めることができる。

- ・出願人・権利者：関連技術を出願している可能性の高い会社から見る
- ・特許分類：関連性の高い特許分類が付与されている特許から見る
- ・日付：開発トレンドを加味し、関連技術が出願されている可能性の高い期間の優先度を上げる
- ・キーワード：特徴語や特定キーワードの存在有無・出現頻度により選別し優先順位を付ける
- ・ファミリー情報：特許発行国数が多いファミリーの優先順位を上げる
- ・引用・被引用情報：引用数・被引用数が多い特許の優先度を上げる
- ・スコア：特許スコアの高い特許の優先度を上げる
- ・独自指標（社内分類など）：社内独自のデータを追加して優先度をつける

クリアランス調査においては、優先度の高い特許文献から順に、重点的に調査していくことが特に重要である。なぜなら、膨大な数の特許検索集合から重要な特許を見つけなければならぬ状況では、調査担当者のモチベーションの維持に非常に効果的に働く¹⁾と共に、検索式の妥当性確認と検索式修正・再検索が早い段階でできれば調査全体にかかる時間が短縮できるからである。また、早期の重要特許発見に繋がるため侵害可能性があれば早期に対策検討に入ることもできる。

重点を置くべき範囲の選定ヒントとして、例えば出願人は、自社製品を販売した際①差止め請求をされる可能性がある企業、②ライセンス

契約を求められる可能性がある企業、③部品・材料の購入先やその競合企業、といった具合に細分化すると重点を置くべき範囲を選びやすくなるといわれている³⁾。

また、重点を置くべき範囲の選定ヒントが分からない場合や他に関連性が高いものを抽出したい場合には、統計分析機能を用いると良い。統計分析機能を標準で備えている特許検索データベースも多く簡単に利用することも可能であり、ダウンロードしたデータで表計算ソフトやその他の特許分析ツールを利用することも可能である。

3. 2 事例紹介

(1) 独自情報の追加

主要な特許検索データベースには、社内分類や社内コメントなど独自情報を保存・共有できる機能が搭載されている。例えばThomson Innovation⁴⁾においては、このような機能として「カスタムフィールド」機能があり、特許情報共有を促進するために必要な社内独自情報を格納することが可能である。カスタムフィールドを利用して特許に関するコメントや社内特許分類、社内特許評価など様々なデータをThomson Innovationの特許データと一緒に格納することで、出願人や特許分類など標準的に取得可能な項目と同じように、独自情報を優先度付けの指標として用いることができ、調査業務の効率化や情報共有が促進される。

Thomson Innovationは、各特許レコード（公報単位または発明単位）に評価情報等独自情報を登録することができる。登録方法としては、エクセルなどのアップロードファイルから一括でアップロードする方法や、個別に登録する方法が存在する。

(2) スコアの活用

特許検索データベースの中には、独自の計算

本文の複製、転載、改変、再配布を禁止します。

式を用いて特許の重要性に応じたスコアを算出する機能を有するものがある。ここでは引き続き、Thomson Innovationを例として紹介する。

Thomson Innovationは「関連性」というスコアを用いることが可能である。「関連性」は、検索キーワードが検索対象フィールドに現れた回数をカウントして算出している。

また、表示設定画面にて、「引用特許数」及び「被引用特許件数」を表示することも可能である（図2）。

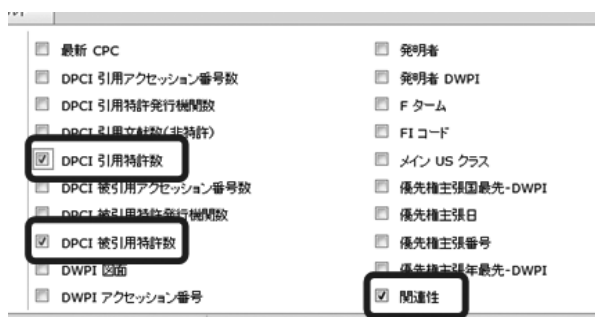


図2 表示と並び替えのオプション

関連性で表示順をソートすることで、関連性が高い特許から読み込むことができる（図3）。

また、表示設定画面で「引用特許数」や「被引用特許数」も表示するように設定することにより、引用・被引用情報を指標とした優先順位付けが可能となる。関連性だけでなく、引用数、

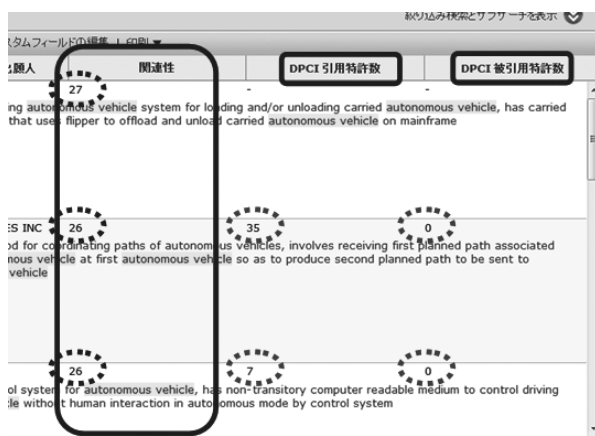


図3 表示順のソート

被引用数を統合して指標化することで優先順位付けの精度が高まる。ここで、引用数、被引用数はエクセルにエクスポート可能であるが、関連性はエクセルにダウンロードすることができないため必要な場合は一覧表示画面からコピー＆ペーストをする。

(3) 特許分析ツールの活用

多数のプレイヤーが犇めき合っている分野においては、特許分析ツールを活用して異なった指標の組合せにより幾つもの特許束に仕分ける方法もある。図4は、PAT EASY Z⁵⁾を用いてマトリクス（縦軸：出願人、横軸：FI技術分類）に特許検索集合を仕分けたものである。

濃い特許束を見出しやすくする工夫の一つとして、SDI等でピックアップした特許案件がマトリクス上のどの特許束に含まれているか判別できるようにするとよい。簡易的には、図4で仕分けた各特許束において、SDIでピックアップした特許案件（「適」のフラグを立てて円の中で色分けしている）がどの程度含まれているか塗り分けるだけで、一目で分かるようになる。



図4 マトリクスマップの活用

4. マニュアルスクリーニング

4.1 概要

マニュアルスクリーニングでは、自動で仕分

けできないものを、目視でさらに仕分ける。例えば、「ログ」と検索した場合の「アナログ」など、短いキーワードを検索に使用すると多数のノイズを含むことになり、こういったキーワードはオートスクリーニングの仕分けにも使いにくい。また、機械的な構造に特徴がある場合でも、図面の画像を用いた絞り込みをオートスクリーニングで行うことは難しい。

本工程では、各特許文献の読み易さを向上させるためのデータ加工が効率化に寄与する。

近年の多くの特許検索データベースには下記のような機能が標準で備わっており、これらを活用して効率よく読込んでいくことが可能である。

- ・ 公報表示画面のカスタマイズ
- ・ カスタムフィールド
- ・ 請求項の読み易さ向上
- ・ ハイライト表示
- ・ 文献並べ替え機能

また、読込みの助けとなる付加情報を加えることにより閲覧をより容易に実施できる環境を整えることも重要となる。具体的には次のような付加情報を加えることが考えられる。

- ・ 機械翻訳データ
- ・ 評価情報反映（評価結果やコメント）
- ・ ファミリー情報
- ・ 引用・被引用情報
- ・ 図面データ
- ・ PDF公報リンク

ここでも、特許検索データベースに備えられた機能だけではスクリーニングの効率化が図れないような場合は特許分析ツールを活用するのも一考に値する。一般的なマップソフトでも査読画面を調査分野に適した表示に適宜カスタマイズできるものであれば活用することができる。

4.2 事例紹介

(1) 特許検索データベースの表示画面のカスタマイズ

カスタムフィールドを活用して独自情報を追加し、特許検索データベースの表示画面をカスタマイズすることにより読み易くする。例えば、Thomson Innovationのカスタムフィールドに対応日本特許の要約や請求項の情報を格納することで、英語情報に加え、格納した日本語情報を参考にして特許内容の閲覧が可能となり読込が大きく効率化される。日本語情報は別の特許検索データベースから一括取得したうえで、エクセルアップロードファイルとして一括登録する。

(2) 特許分析ツールを活用した表示画面のカスタマイズ

請求項の数が従属項も含めて多数あると、どれが独立項でどれが従属項かを選別するだけでも一手間である。独立項だけを表示する機能を搭載する特許検索データベースとして、TotalPatent⁶⁾がある。

また、Espacenet⁷⁾は無料のクレームツリー表示機能を公開しているので利用するのもよい。マップソフト等を活用すれば、クレームツ

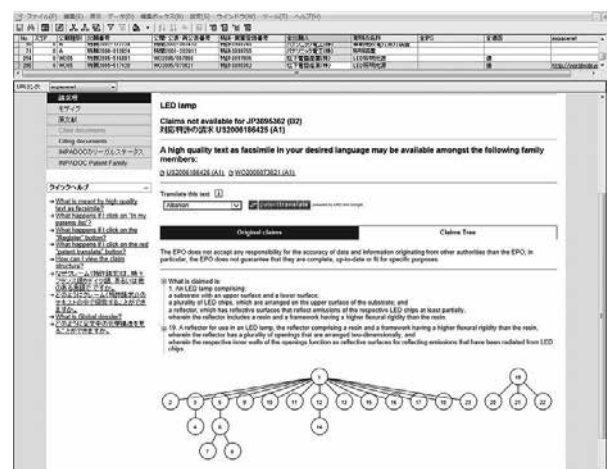


図5 パテントマップEXZの表示にEspacenetクレームツリーを組み込んだ例

リー等のようにスクリーニングに必要な情報を複数の情報源から一画面に集約することができ、効率的な判断プロセスをサポートする。

図5は、パテントマップEXZ⁸⁾を活用して、4.1で記載した付加情報に加えEspacenetのクレームツリーを一画面にカスタマイズしたものである。確認工数の削減もマニュアルスクリーニングにおいて効率化に繋がる。

(3) ハイライト位置の表示機能

多くの特許検索データベースには、明細書中の任意の文字をハイライト表示させる機能が備わっており、注目したいキーワードが見つげやすくなっている。また、いくつかの特許検索データベースには、特許文献中のどの辺りにハイライトされたキーワードがあるかを表示する機能が備わっており、これを利用すれば、特許文献を読み進めていく上で、問題特許かどうかを確認するための文章を早急に見つけるのに役立つ。図6はPatentSQUARE⁹⁾のヒットバーの表示画面である。

前述の通り、クリアランス調査では再現性が



図6 PatentSQUAREのヒットバー

重要視されるため、検索式でNOT検索（特定キーワードによる除外）の使用は最小限に止めるべきである。例えば「A」に関する特許を拾いたいAの下位概念「B」に関する特許は不要な場合、「A」で機械検索して得られた各特許について、全独立請求項に「B」が必須要件として記載されているか否かを目視確認する必要がある。仮にこれを機械検索の段階で「A NOT B」と論理演算すると、「B」が従属項にしか記載されていない特許や「B」が任意要件としてしか記載されていない特許も検索結果から除外されてしまう。そのため、特定キーワード「B」が独立クレームに含まれるか下位クレームに含まれるか一目で確認できればクレームを通読せずにノイズか否か判断できる。

(4) 図面の視認性向上

調査対象の技術分野や出願国によっては、特定の図面が存在するか否かによる優先度を付けられる場合がある。ここでいう「特定の図面」とは、例えばクリアランス調査の対象がユーザーインターフェースそのものである場合には、明細書には画面の表示状態、あるいは最低でもモニターが図示されているのが一般的である。これを考慮し、レイアウト変更で複数図面を一度に表示することで、明細書を査読することなく優先度を付けることができる。図7は、NewCSS/JP-NET¹⁰⁾で複数の図面を一画面に表示させた例である。

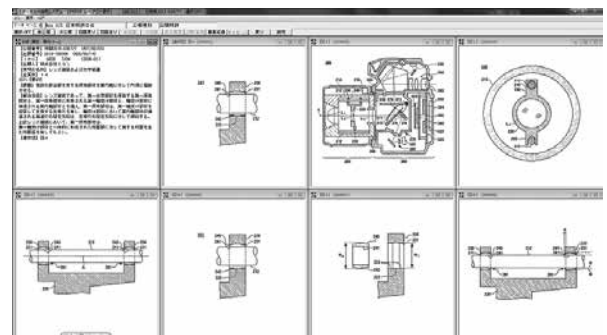


図7 NewCSS/JP-NETのレイアウト表示機能

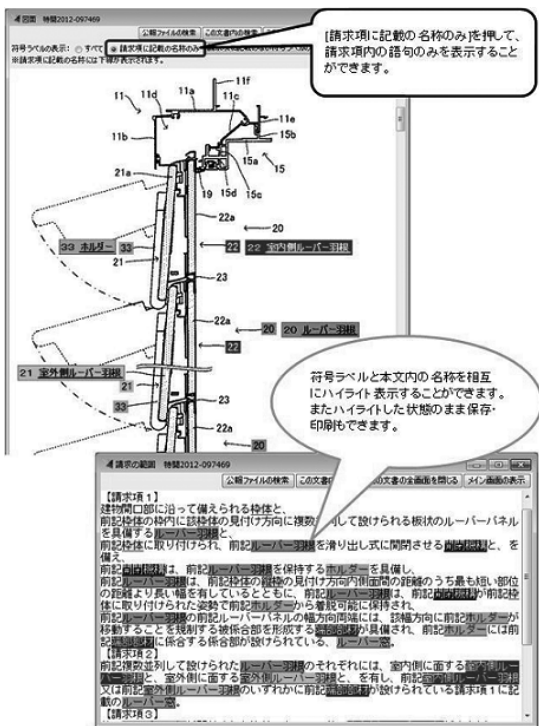


図8 図面に構成要件を記載 (公式サイト¹¹⁾ より引用)

また、特許ストーリー 公報読解支援ソフトウェア¹¹⁾を使うと、図面の各符号に対応する構成要件を自動的に図面に記載することができる (図8) のので、このツールを使うことで図面を理解する労力を少なくし時間を短縮することができる。

5. 重要な特許の抽出後

スクリーニングにより濃い集合を抽出した後は、各案件を詳細に読み込んでいく。4章で述べた表示画面のカスタマイズなどを行っておくと公報が読みやすくなり、理解が深まると共に時間の短縮にもつながる。

また、問題特許が見つかった際には、引き続き、その問題特許に無効理由がないか、或いは、その特許の効力が及ばない先行技術を探すのが普通である。また、抽出した問題特許に基づいて、他に類似する内容の特許がないか、スクリーニングの漏れがないか再調査することも行うであろう。

ここでは、再度検索式を立てる、という議論をするのではなく、明らかとなった問題特許から、他に類似特許、関連特許がないか、簡易に確認する手法を紹介する。

(1) 引用・被引用の確認

多くの特許検索データベースにサイテーション機能が搭載されている。引用・被引用は、特許庁の審査官が「近い技術」と判断した特許文献の紐付けであり、紐付けられた各特許文献の近さについては、ある程度の信頼性があると考えていだろう。よって、サイテーション機能を用いて紐付けられた引用・被引用特許を確認することで、問題特許に近い類似特許にヒットする確率は比較的高いと考える。また、引用文献を確認することで、問題特許の権利範囲がより明確になり、回避案の検討材料としても活用ができる。

2014年に日本でもサービスが始まったAmberScope¹²⁾という特許分析ツールを用いれば、DOCDBに収録されている各国特許文献の引用・被引用情報を利用して、よりグラフィカルに引用・被引用のネットワークを視認することができる。

「AmberScope」は、独自のアルゴリズムによって、問題特許に関連する特許の重要度を円

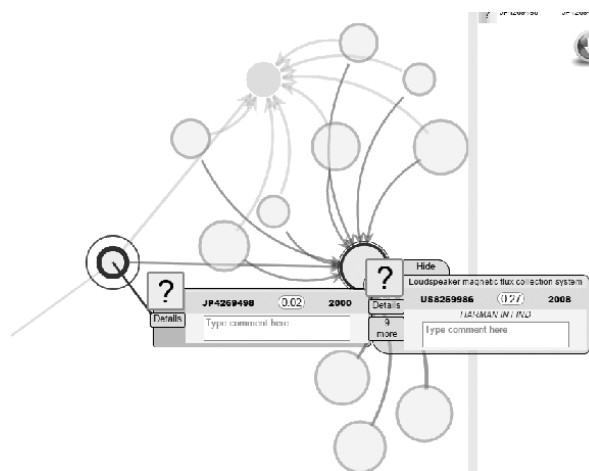


図9 AmberScopeの分析結果

の大きさで表現し、関係性の強弱を線の太さで表現するので（図9）、引用・被引用の特許文献の中でも、より確認が必要と思われる重要度の高い特許を素早く見つけることができる。

(2) 概念検索機能の活用

概念検索機能もサイテーション機能と同様、主要な特許検索データベースに標準装備されている。既に問題特許が判明して、その問題特許の類似特許を探す手段として、ここでは、HYPAT-i¹³⁾の概念検索機能を活用した例を紹介する。

概念検索では、通常、検索文字入力フィールドに検索したい対象を表現した文章を入力して検索するのが一般的である。ここに問題特許の請求項の文章若しくは、要約文章を入力して検索を行う。

その結果、概念検索のアルゴリズムが自動的に判断した類似特許群が抽出され、類似度の高い順番にリスト表示する。

これだけで問題特許に近い特許が見つかることもあるが、条件を追加してより近いものが上位に上がるようにすることもできる。

HYPAT-iの概念検索には、「条件変更機能」が備わっており、概念検索のアルゴリズムが自動的に切り出したキーワードに重み付けをした

り、新たなキーワードを追加するということができる。図10は、その設定画面の一例である。

更に、「絞り込み機能」を利用して、問題特許から抽出したFI、出願人名、発明者名で絞り込むことで、より類似度の高い特許文献を抽出することもできる。

6. その他ツールの紹介

スクリーニング効率化を目的に開発された機能について紹介する。これらは特定のシステム・ツールにのみ備わった特徴機能であり、誰もがすぐに使えるようなものではないが、その分強力な手段といえる。スクリーニングに着目して実装された機能は未だ少なく、今後の発展に期待したい。

(1) 「THE調査力_クラウド¹⁴⁾」のハイライトソート

THE調査力_クラウドに備わる機能には、重要特許には必ず含まれるであろうキーワードを設定すると、そのキーワードが多く含まれる順に母集団の中の特許文献を並び替える。通常のハイライトではキーワードがあるかないか、又はどこにあるかくらいしか判断できないものを、この機能により、ハイライトの数でランキングをつけることができる。

本機能により、設定したキーワードが多く含まれる特許文献がリストの上位に集まるので、「濃い集合から見るメリット」と同等の効果を得られる。

(2) クレームマップ及びクレームクラスター

インパテック株式会社が提供する「クレームマップ¹⁵⁾」は、特許請求項の分析に特化した支援ソフトで、全請求項のツリー表示やキーワードハイライト等の他に、主従関係にある請求項間のキーワードの差分を明示する機能が実装されている（図11）。従属項で付加された構成要

重要度	検索語
<input type="radio"/> ▾	首
<input type="radio"/> ▾	指向
<input type="radio"/> ▾	筐体
<input checked="" type="radio"/> ▾	スピーカ
<input type="radio"/> ▾	音声
<input type="radio"/> ▾	端末
<input type="radio"/> ▾	上端
<input type="radio"/> ▾	装置
<input checked="" type="radio"/> ▾	上方

図10 HYPAT-iの条件設定画面

素のキーワードを一目で把握でき、クリアランス調査時の査読効率が向上する。



図11 クレームマップの詳細画面

また、同社が2015年4月にリリース予定(2015年3月現在)の「パットクレームクラスター¹⁶⁾」(図12)は、大量の特許検索集合をユーザが事前に作成した辞書を基に自動仕分けするクリアランス調査支援ツールである。



図12 パットクレームクラスターの詳細画面

自動仕分けされた特許案には辞書に基づくフォルダ名がつけられて関連公報を格納する。調査担当者は、より優先度の高いフォルダ(特許案)から調査を開始したり、担当者間でフォルダを分担したりすることも容易になる。格納された各公報には、最終請求項の末尾に記載された「…を特徴とする○○。」「…から成る○○。」の○○部分に該当するキーワードを“発

明カテゴリー”として各請求項に付与され、査読工程を支援する。

クレームマップとパットクレームクラスターとは、別個のツールの為、両者を併用することにより調査効率は一層向上する。ユーザの利便性を考慮すると将来的にはクレームマップにパットクレームクラスターの機能が搭載されることを期待したい。

7. 文書管理上の注意事項

スクリーニング作業を終えた後、スクリーニング結果を共有するための資料を作成する場合がある。ここまで、クリアランス調査を題材にスクリーニングについて述べてきたが、クリアランス調査を通して作成する資料は、後日米国で訴訟が生じる可能性を想定して、他社権利侵害を自認するような表現とならないよう十分に配慮する必要がある¹⁷⁾ことを注意しておきたい。

8. おわりに

大量の特許を、ただ闇雲に読んで理解することは大きな労力となる。優先度を付けて読み込む方針を立てること、また、データ加工や表示画面のカスタマイズにより読み易い環境を整えることが、この労力の低減に繋がる。スクリーニングを目的として開発された機能やそれを備えたツール類は数が少なくあまり一般的ではないので、各自が使えるもので工夫することが重要である。

本稿で紹介したような手法をもとに、特許検索データベースや特許分析ツールを活用することにより、少しでもスクリーニングの労力を減らせることができれば幸いである。

なお、本稿は、2014年度情報検索委員会第1小委員会第3WGメンバーである足立圭(協和発酵キリン)、小田将文(TOA)、鎌田翔一(シャープ)、鈴木浩三郎(ミットヨ)、高岡恵理(ライオン)、米須秀樹(旭硝子)によるものである。

注 記

- 1) 知的財産情報検索委員会第3小委員会, 特許情報検索における教育に関する提言—スクリーニング手法を中心に—, 知財管理, Vol.61, No.10, pp.1523~1535 (2011)
- 2) 鈴木利之, 効率的で漏れのない特許調査 (後編), 知財管理, Vol.60, No.2, pp.303~308 (2010)
- 3) 酒井美里, 知財部員のための侵害予防調査—検索着手前に考慮すべき点と検索式作成のテクニック—, 知財管理, Vol.65, No.2, pp.266~271 (2015)
- 4) Thomson Reuters : Thomson Innovation
<http://ip-science.thomsonreuters.jp/products/ti/> (参照日2015/3/27)
- 5) インパテック株式会社 : PAT EASY Z
<http://www.inpatec.co.jp/software/patentmapeazy-z> (参照日2015/3/27)
- 6) LexisNexis : TotalPatent
<http://www.lexisnexis.jp/ja-jp/Products/totalpatent.page> (参照日2015/3/27)
- 7) Espacenet
http://worldwide.espacenet.com/?locale=jp_EP (参照日2015/3/27)
- 8) インパテック株式会社 : パテントマップEXZ
<http://www.inpatec.co.jp/software/patentmap> (参照日2015/3/27)
- 9) パナソニック株式会社 : PatentSQUARE
<http://panasonic.biz/it/sol/patent/> (参照日2015/3/27)
- 10) 日本パテントデータサービス株式会社
New CSS
<http://www.jpds.co.jp/jp-net/newcss.html> (参照日2015/3/27)
JP-NET
<http://www.jpds.co.jp/jp-net/jp-net.html> (参照日2015/3/27)
- 11) アイビーリサーチ株式会社 : 「特許ストーリー」
公報読解支援ソフト アナライザー版
<http://www.ibr.co.jp/lessonpat/analyzer.html> (参照日2015/3/27)
- 12) Ambercite日本公式サイト
<http://ambercite.jp/> (参照日2015/3/27)
- 13) 株式会社発明通信社 : HYPAT-i
http://www.hatsumei.co.jp/hypat_i/ (参照日2015/4/1)
- 14) アイ・ピー・ファイン株式会社 : THE調査力クラウド
http://www.ipfine.com/TIP_Cloud/ (参照日2015/3/27)
- 15) インパテック株式会社 : クレームマップ
<http://www.inpatec.co.jp/software/claim> (参照日2015/3/27)
- 16) インパテック株式会社 : パットクレームクラスター
<http://www.inpatec.co.jp/software/patclaimcluster> (参照日2015/4/15)
- 17) ヘンリー幸田, 日米特許紛争スーパーマニユアル—特許法務 : 戦略と予防, 発明協会 (1992)

(原稿受領日 2015年4月20日)