

機械学習や自然言語処理技術を用いた知財リーガルテック^{*}について



アイ・ピー・ファイン株式会社 取締役副社長 平尾 啓

要 約

日本の研究開発の質・量ともに向上させることが急務とされる中、研究開発段階において研究開発部門への弁理士の支援・提案が期待される。当該技術知識においては研究開発者に劣るのは否めないが、弁理士には研究開発者とのコミュニケーションにより研究開発の質・量を向上させ発明をより高次に持ち上げる力量があるのである。ただし、会話をするだけではお互いの理解・納得感は醸成しにくい。そこで、適切な知財リーガルテックを活用し、議論の趣旨が見える化することが必要となる。現在、非常に多くの知財リーガルテックが販売されているが、研究開発者とのコミュニケーションに有効な知財リーガルテックを弁理士自らが選択し、実際に使用する必要がある。例として、研究開発力向上のために、研究開発者レベル向上とコミュニケーションのスピードにあった図表の簡易作成ツールを、調査業務の効率化では、調査の目的、内容によって使い分ける知財リーガルテック選択のポイントを説明した。

目次

- 1. はじめに
- 2. 研究開発力強化への知財リーガルテックについて
 - 2. 1 特許出願数の年推移の現状
 - 2. 2 発明の質・量の向上への取組
 - 2. 3 開発戦略策定のための知財リーガルテック
 - (1) 3つの目
 - (2) 社内分類（独自分類）の重要性
 - 1) SDB（Strategy Data Base）について
 - 2) 特許の群管理の重要性
 - (3) 知財情報分析の課題
 - (4) 社内分類作成、見える化支援のための知財リーガルテック
 - 1) 社内分類付与及び特許の紐づけツール：例「The Tree」
 - 2) 社内分類等簡易マッピングツール：例「SPM」
 - (5) 知財分析支援のための知財リーガルテック
 - 1) ツールを開いたら時系列マップ
 - 2) 図表作成が手軽、図表がコミュニケーションツール
 - 2. 4 第2章のまとめ
- 3. 調査業務効率化への知財リーガルテックについて
 - 3. 1 調査業務に求められること
 - 3. 2 調査者のニーズへの知財リーガルテックについて
 - (1) 『ハイライト』機能：例「THE 調査力^{*}」
 - (2) 『ハイライトソート』機能：例「THE 調査力」
 - (3) 『ハイライト折込み』機能：例「THE 調査力」
 - (4) 文書の類似順ソート機能：例「Deskbee^{*}」(人工知能)
 - 1) 「Deskbee」(人工知能ツール)の基本機能説明
 - 2) 「Deskbee」の結果表示機能
 - 3) 「Deskbee」の初期的課題
 - 4) 「Deskbee」の課題解決
 - 5) 「Deskbee」の特長
 - 3. 3 知財調査効率化のための知財リーガルテックについて
- 4. まとめ

1. はじめに

本著作を依頼されるにあたり、何人もの弁理士の先生方にインタビューさせていただいた。弁理士に必要な知財リーガルテックはどの範囲であるのかがわからなかったからである。調査業務そのものをしない弁理士事務所がある一方、調査分析に非常に熱心な弁理士事務所がある。クライアント（企業内弁理士は自社）の研究開発力の強化に向けた活動が弁理士にとってとても大切な分野となると多くの先生にインタビューで教えていただいた。そこで、研究開発者に的を絞ることとし、2章では「研究開発力強化」、3章では研究開発者が大いに負担に感じている「調査業務」の効率化について書かせていただく。

^{*} 登録商標

2. 研究開発力強化への知財リーガルテックについて

2.1 特許出願数の年推移の現状

日本の特許出願年推移⁽¹⁾を図1に示す。2000年の特許出願数：436,865件から2019年：307,969件と緩やかではあるが、減少していることがわかる。この20年では30%の減少である。

2.2 発明の質・量の向上への取組

「不完全な特許は無駄であり、量より質の向上を図るべき」との議論もあり、特許出願数の減少はさして大きな問題ではないと言う方々もおられる。しかし、中国をはじめ特許出願数は世界的に増加傾向にあり、顕著な減少傾向にある日本との違いも見ることができる。

量か質かの議論の中で、3つの話を紹介する。

① 「改善提案制度」の事例

改善提案制度を導入している企業のある工場での話で、工場長から年初に「改善提案制度の質的改善をする。今後、低レベルの改善提案は受けつけない」と方針発表されたところ、改善提案数が激減した。高レベルの改善提案だけになったとしばらく運用したところ、ついには高レベルの改善提案も大幅に減少したという事例があった。

② 「ハインリッヒの法則」から学ぶもの

ハインリッヒは、アメリカの損害保険会社にて技術・調査部の副部長をしていた1929年に「ハインリッヒの法則」を発表した。1件の大事故・災害の裏には、29件の軽微な事故・災害、そして300件のヒヤリ・ハット（事故には至らなかったもののヒヤリとした、ハットした事例）があるとされるものである。

軽重事件の構造はピラミッド構造となっており、すそ野の軽い事件が少なくなれば上も高くなる（重い事件が減る）と言える。

③ 「研究開発者の知財経験・関心」の視点

筆者経験からの話であるが、ある会社において、座学での知財教育を知的財産部門が研究開発部門に対し熱心に行っていたが、なかなか研究開発者の知財への関心は高まらなかった。そこで、特許出願作業にしっかり研究開発者を巻き込む仕組みとしたところ、研究開発者の知財への関心が一気に高まった。特許調査においてもしっかり特許を読み込んだり、知的財産部門とのコミュニケーションも熱心に行うようになった。個人的な問題もあると思うが、人は経験していない物事には慎重になったり、関心が薄かったりすることが多いようである。この仕組みも特許出願が前提となっており、そもそも特許出願の数が減るとこのような経験ができる研究開発者が減るとのことである。

要は「量か質か」の二者択一ではないのではない。ただ、「質」を追い求めると「量」が減り、そのうえで「質」が下がる。では、低レベルの特許を出願し続けるのを容認するのかということ、そうではない。

低レベルの発明を切り、上澄みを出願するという単純な従来の方法ではなく、質的レベルは関係なく、いかに全体に劇的なレベルアップを図るかが重要になってくるのではないと思われる。

それには各会社の努力も必要であるが、弁理士の役割が大きいのではないと思う。開発戦略策定時、研究開発テーマ策定時、研究開発時、発明完成時、出願作業及びその後とある中で、一般に、弁理士の活動は出願作業時及びその後のようである。私の知り合いの弁理士には、研究開発のより上流工程で知的財産部門

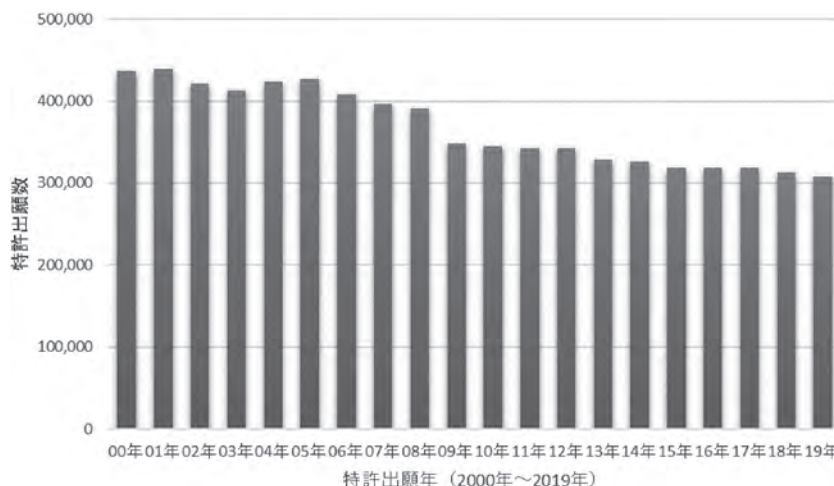


図1：日本の特許出願数の年推移

だけではなく研究開発部門と熱心にコミュニケーションをされている方が多くおられる。発明の質量両面のレベルアップを図るためには、弁理士の活躍が期待される。

弁理士のコミュニケーションの武器として、知財知識・経験だけではなく、知財リーガルテックが欠かせないものではないか。

2. 3 開発戦略策定のための知財リーガルテック

(1) 3つの目

一般に経営、マーケティングの戦略策定において「3つの目」は重要である。

表1に示すように虫の目、魚の目、鳥の目が「3つの目」である。

通常の特許調査は「虫の目」で、一つ一つを読み込み、問題特許・重要特許などを見つける作業である。しかし、戦略構築上に必要な目である「魚の目」「鳥の目」は十分に使えていない。開発戦略策定のためには、この3つの目を効率よく組合せ利用できる知財リーガルテックが必要である。

表1：3つの目

3つの目	視点	調査実務
虫の目	個々情報の評価	個々を読み込み問題特許、重要特許を発見
魚の目	時流・トレンドを見る目	年推移などから競合・技術動向など
鳥の目	全体を俯瞰する目	全体像から自他社強み弱み分析など

(2) 社内分類（独自分類）の重要性

1) SDB (Strategy Data Base)⁽²⁾⁽³⁾について

開発戦略策定において重要なことは、社内の関係者が十分な議論をすることができることである。そのためには社内で「皆が理解できる言葉」により資料が作成されていること（見える化が図られていること）が必要である。

この「皆が理解できる言葉」については、鶴見隆⁽²⁾中村栄⁽³⁾がSDBを提唱している。SDBは鶴見隆によると、『自社の技術分類、戦略的に重要な判断を個々特許案件に付したものとされる。一般に「社内分類（独自分類）」と言われるものである。以下の記述は「社内分類」とする。

2) 特許の群管理の重要性⁽⁴⁾

「技術の複合化が進んでいる分野においては、一つの商品を数百にも及ぶ特許権で保護することもあり、そのような商品を扱う企業においては、研究開発成果である発明を個々に単体でとらえるのではなく、商品や技術テーマ等との関係で「群」としてとらえていく

ことの必要性に迫られている」とあるように複数の特許を塊として扱う必要がある。そのためにはその塊に名前『社内分類』が必要である。

また、当報告書には「課題×解決手段」のグラフ作成が開発戦略構築のためには非常に有効であるとあり、「課題」「解決手段」それぞれに社内分類が付されることが非常に社内コミュニケーション、理解・納得を得るのに重要であることがわかる。

(3) 知財情報分析の課題

上述の3つの目、社内分類の重要性などは10年以上前に提唱されたものである。最近では「IPランドスケープ」などと言われ、過去には「経営に資する知財分析」など、様々な呼称でその重要性は議論されてきている。講演会、講習会など多くの場で取り上げられて関心も非常に高い。しかし、なかなか多くの会社からの成功事例が聞こえてこない。多くの原因があると思うが、ここでは図表を作成する知財リーガルテックに関係するものを1点あげる。

① 研究開発戦略・研究開発テーマ決定（提案）時において

特にリーダー・研究開発所長など上位者に対する知財分析報告時に多く聞かれることであるが、説明終了時に上司から「でっ？何？！」と問われることが多いとのこと。これは知的財産部門から研究開発部門に対するプレゼンテーションでも聞かれることが多い。「図表はきれい、趣旨不明」とも。そこで報告者は続く言葉として、「こういう傾向がみられます」と、すると先方から「傾向はいいけど、私たちに何をしろというの?」。そこで言葉に詰まることが多いようである。その場を取り繕い、改善版の作成だけではなく、出席者のスケジュール調整もし、やっと2週間後に改善版を持参すると、「何の話だった? あっ! それか。それは問題があつて変更したよ。もういいよ」と。

② 研究開発テーマ検討時などの議論をする場において

このような場面においても、知財情報分析結果の利用は非常に大切であるが、①と同様、議論の場で臨機応変に図表を作成し、議論を深めていく取り組みはまだ十分ではないようである。図表の作成が2週間後ではなく次の日と早くなったとしても議論は長時間かかり低調となりがちである。

①においては、「分析担当者が仮説検証型の思考で

はないこと」「知財情報分析作業への期待のずれ」が原因のようであるが、①②に共通のものは、知財リーガルテックの使い勝手が原因のようである。つまり「コミュニケーションの場で見たい図表を作成できる」というツールの簡便さが求められるということ。

ツール提供者側も使い勝手の改善について取り組んではいるが、お客様の多様なご要望をツールに反映すると、どうしても機能が多くなりかつ複雑化するのは避けられない。そのためにツール説明会など頻繁に開催しているが、ツール利用者はついていけないのが現状のようである。また費用の面からも、高機能ツールの導入、ましてや複数導入は二の足を踏むケースもあるようである。

以上から、今後は「分析はこのツールを使用するとなんでもできます」というオールインワンツールではなく、ある業務に特化したツール、そしてそのツール同士をつなぐハブ機能を持ったツールを使い分けることも必要ではないかと思う。必要な機能のみのツール導入であれば費用面からもリーズナブルであり、使用方法の習得も楽であると思われる。

私の知り合いの弁理士からは「まずは研究開発部門との発明相談でのコミュニケーションの高度化のためのツール」「研究開発の進め方についての支援・提案をする場が多くなったので実務的なツール」が欲しいとの声を聞くことがある。この部分に限定・特化した機能をいくつか紹介したい。

(4) 社内分類作成、見える化支援のための知財リーガルテック

一般に「社内分類」付与作業は非常に手間がかかり、また緻密に「社内分類」を作成すればするほど、メンテナンス作業が困難であるとも言われている。よって付与作業、メンテナンス作業の手間により、「社内分類」作成の重要性・必要性は実感しながらも、二の足を踏むことや、せっかく「社内分類」作成したがメンテナンスできずお蔵入りするケースが散見されている。この課題についていくつかのツール提案がされているので紹介する。

1) 社内分類付与及び特許の紐づけツール：例『The Tree』

社内分類付与及び特許の紐づけ作業について、一般に次のような方法がある。

① 複数人が KJ 法、UMethod⁽⁵⁾などで広いテーブ

ル、壁、ホワイトボードなどに付箋などを張りながら作成する。付箋なので試行錯誤で張り替えながら簡易に完成に近づけることができる。また画面が大きいので一望一覽しやすい。ところがその後、電子化し二次利用するのに難しい時がある。また結果の保存やメンテナンス作業が困難なことがある。

② Excel（登録商標）などに入力しながら作成する。電子データなので完成後はハンドリングしやすいが、作成時は画面が小さいので一望一覽を参加者全員で行うのに制約がある。また、Excel 内の社内分類、特許紐づけ作業は、数が多い場合、特に項目の移動が困難となり、またメンテナンスが困難なことがある。

このような作業をより効率的に行うためにパソコン上で、大人数で試行錯誤する中で分類・特許紐づけ作業をドラッグ&ドロップで簡易にでき、完成後もプリントアウト、またメンテナンス作業も非常に簡易にできる機能が必要になる。

一例として、「The Tree」を紹介する。

特長は、

- ① 社内分類をこの場で書き込むことができ、その分類の階層間での移動がドラッグ&ドロップで簡単。また分類数、階層数などに実務レベルにおいて制限がないこと。
- ② 分類に特許を紐づけることもドラッグ&ドロップで簡単。
- ③ 必要時に該当する特許文書を簡単に確認することができる。
- ④ プリントアウトすることが容易にできる。
- ⑤ ツール内だけではなく Excel での保存が簡単で、保存したものを簡単に呼び出し、継続的にメンテナンスができる。
- ⑥ 「THE 調査力」（後述）をハブツールとして、他のツールとの連携が可能である。他のツールで作成した分類の取り込みも可能であること。

2) 社内分類等簡易マッピングツール：例『SPM』

マッピングツールは社内分類と出願人、発明者、出願日などの日付などを組み合わせた図表などを簡単に作成することができるが求められる。その時に必要な条件は以下の通りである。

- ① 横軸、縦軸となる項目を簡単に選択できること。変更も簡単であること。作成した図表の保存が簡易にできること。

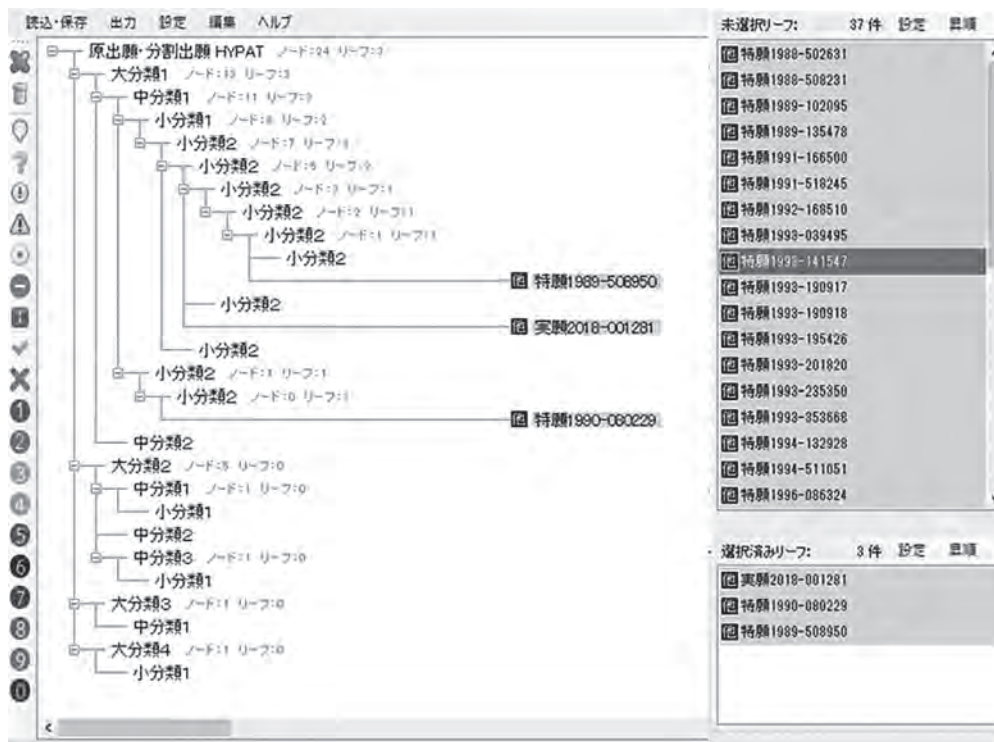


図 2：ツール「The Tree」の画面



図 3：ツール「SPM」の画面

- ② 軸内の項目の表示順の変更が簡易にできること。
- ③ 図表内の社内分類の階層を簡易に変更できること。
- ④ 図表上で気になった部分（例えば、棒グラフの棒、泡グラフの泡のようなグラフのパーツ）をクリックすると、それに該当する特許が簡易に表示されること。
- ⑤ 図表の表示速度が、瞬く間であること。
- ⑥ 他のマッピングツール、Excel などからのツール乗り換えが簡単にできること。

特に、例えばある会社に対し第三者がマッピングツールを使用し特許分析などのアドバイスをする場合、相手（クライアント）の会社が契約するマッピングツールをすべて契約し、それらを用いてアドバイスするにはスキル上、経済上困難である。ただし、アドバイスする側が自分の得意なマッピングツールを最低一つ使用しながら、相手（クライアント）に例示できることが非常に重要である。これには上記⑥のようにデータの簡単な移動が重要な鍵となる。

①～⑥ができる、マッピングツールの事例として、『SPM』を紹介する。

よく利用される、棒グラフ、円グラフ、泡グラフは当然作成できるうえ、様々な戦略検討時に必要と思われるマップ作成機能がある。

（5） 知財分析支援のための知財リーガルテック

上記に「3つの目」の紹介をした。現在、虫の目である日常調査業務と、魚の目、鳥の目である知財分析業務は別々に扱われることが多い、特に研究開発に携わる研究開発員にとって、知財はひたすら1件ずつ読み続けるものであり、なかなか魚の目、鳥の目での視点からのアプローチはしにくいものである。多数の難解な特許文書、かつ重要、不要の特許が玉石混合する集合をひたすら注意力をもって読み続けるという難行苦行をする中で、研究開発に関わる情報として非常に重要な知財情報を嫌う傾向が出てきていると聞くことが多くある。

そこで虫の目に加え、魚、鳥の視点となるものを組み込んだら、研究開発者は競合他社、業界の動向をチェックしながら重要な特許から読んでいくことができるのではないかと、盲目的に特許を読み込む単なる「作業」から、研究開発テーマの発想や、研究開発業務の一環としての業務になるのではないかとと思われる。

一例として、ツール『THE 調査力』を紹介する。

1) ツールを開いたら時系列マップ

THE 調査力で自分の担当する調査タイトルを開くと、図4の時系列表が表示される。デフォルトとして縦項目に出願人名とすると、出願人毎の特許出願時系列マップとなる。研究開発者は最初に魚の目としてトレンドを目にすることができる。また、自分の競合他社、次に上流の原料資材会社、その次に一般会社などの優先順位をもって、特許調査をすることができる。

また、特許分析ツールを使う上での泣き所は、「作成した瞬間からその図表は古くなる」ということである。つまり、特許は次々と公開されてくることから、その情報を加味した図表を作成しなければ最新にならないということ。しかし、今まで述べてきた通り、一般の特許分析ツールを使用するのは手間がかかるので常に最新にしておくのは困難である。

THE 調査力は、SDI さえ設定しておけば、常に最新の時系列マップがツールを開くと表示される。いちいちマップを作成する手間がない。このように常に最新の分析情報に接することができるのが研究開発者にとって重要ではないかと思われる。

2) 図表作成が手軽、図表がコミュニケーションツール

魚の目であるトレンド表示だけではなく、鳥の目としての鳥瞰する図表についても簡易に作成することが必要である。この図表はコミュニケーションの材料としてのツールとして位置付けることで、きれいな図表であることは必ずしも求められない。必要な機能は、コミュニケーションに合わせた「作成の簡便さ」「表示の速度」である。またコミュニケーションの中で出てくる気になる特許の確認のための「図表から個別案件を表示させる機能」である。

2. 4 第2章のまとめ

以上のように研究開発の成果としての発明の質・量の向上には知財の専門家である弁理士の研究開発部門への支援・助言がなくてはならないものである。しかも、発明の届出後ではなく、もっと開発工程の上流において研究開発者とのコミュニケーションが重要であ



図4：ツール「THE 調査力」の画面1

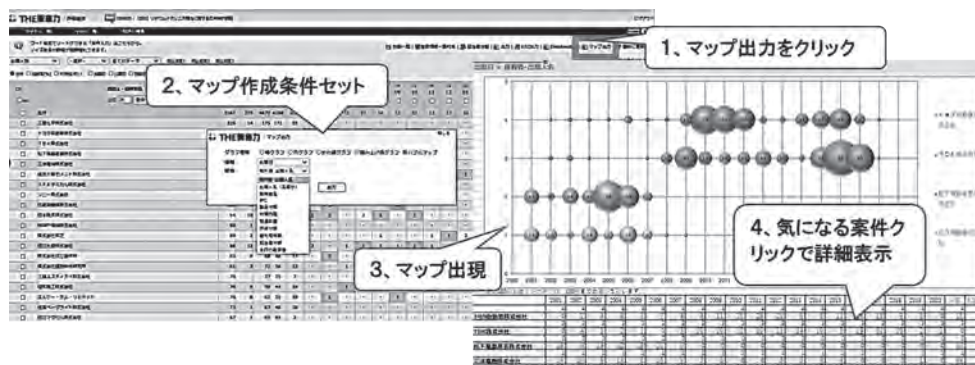


図5：ツール「THE 調査力」の画面2

る。弁理士は技術の詳細については研究開発者に及ばないが知財の専門家である。弁理士、知的財産部門、研究開発部門のコミュニケーションを実りあるものにし、互いの知識、考え方を融合するために、適切な知財リーガルテックを理解し使用できることもこれからは大切である。

3. 調査業務効率化への知財リーガルテックについて

3.1 調査業務に求められること

特許調査において基本的に求められる当たり前の機能として、

- ① 検索結果に漏れがないこと。
- ② できるだけ広い情報が取れること（例：海外データなど収録範囲拡大）
- ③ 検索範囲をよりの確に、絞り込むことができる検索コマンド
- ④ 多数の出力特許数を可能とする（出力項目の選択）。
- ⑤ 操作ヘルプの充実とわかりやすさ

があるが、これらの機能は多くの特許調査データベースにおいて実現されてきて大きな差異がなくなっているので、次のようなより高次の機能が求められるようになってきた。

- ① 読みたいものだけを読む機能
 - ② 読みたいものから読む機能（優先順位付け）
 - ③ 読みたくないものを除く機能
 - ④ 見落としをなくす機能（より広い母集団での調査）
- これらの実現のため、知財リーガルテック例を紹介する。

3.2 調査者のニーズへの知財リーガルテックについて

(1) 『ハイライト』機能：例「THE 調査力」

調査作業は、「特許文書中の『キーワード』およびその組合せを発見し、その重要性を見極める作業」であるともいえる。であるならば、いかに早く『キーワード』を見つけるかについて、ハイライト（ハイライトバー）機能が提供されている。ハイライト表示のされているところを読む（ハイライトされていないところは読まない）。またハイライトが付くことで重要特許の見落としリスクを軽減することができる。

(2) 『ハイライトソート』機能：例「THE 調査力」

ハイライト機能については、確かに上記のような利点がある。しかし、検索結果の出力データを順番に玉石混合状態で見ていくことに変わりなく、この点での

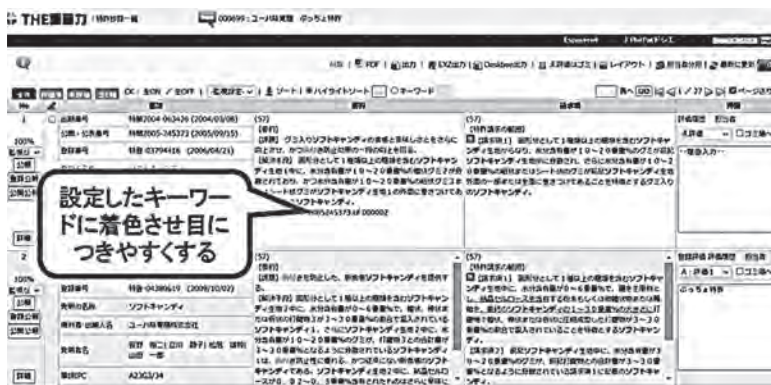


図6：「ハイライト」の画面例

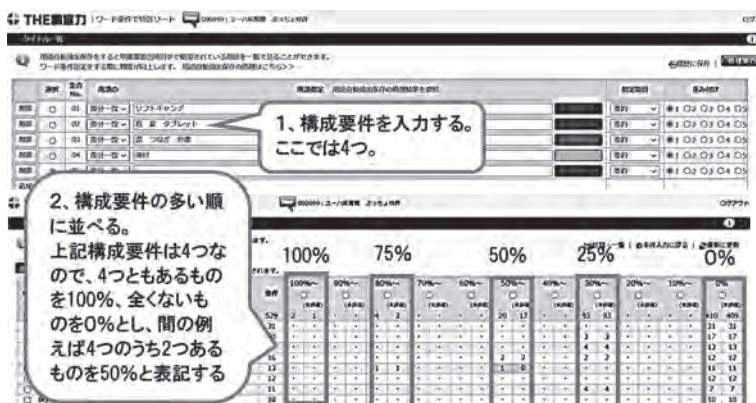


図7：「ハイライトソート」の画面例

調査業務効率化は図りにくい。

重要な特許はキーワードの有り無しでの読み込みとなるが、単独で存在することは少なく、多くの場合、キーワードの組合せでの読み込みとなる。つまり構成要件に含まれるキーワード群の多い順に優先順位をつけて読み込むことが見たいものから見るといことになる。

図7を例とすると、構成要件を入力する（ANDは縦に列挙、ORは横に列挙（間にスペース））。今回は構成要件を4つとする。4つともある特許群は100%と表示する。4つの中の3つある特許群は75%と表示する。順に、4つのうち2つある特許群は50%、4つのうちに1つある特許群は25%、最後に4つのうち一つも無い特許群は0%と表示する。

一般に構成要件の多い順（100%から0%に向かって）に読み込むことで見たいものから読む（優先順位をつける）という機能が実現される。

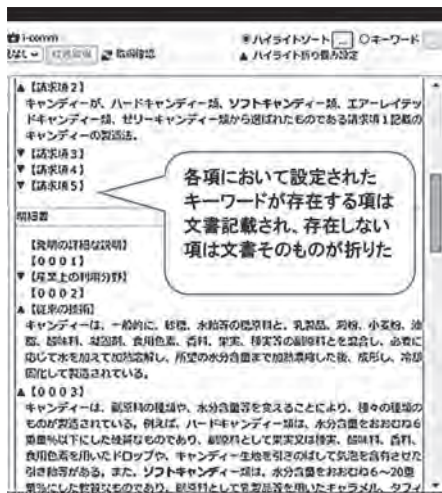


図8：「ハイライト折込み」の画面例

(3) 『ハイライト折込み』機能：例「THE 調査力」

ハイライト機能にプラスする機能として、『ハイライト折込み』機能がある。これは、よりハイライトさせたキーワードをより見やすく、見落としなくするための機能である。

図8に例として、THE 調査力のハイライト折込み機能を示す。複数のキーワードにハイライト表示をさせて表示させた場合、明細書の読み込み時など長文を上から下まで読み飛ばさないように慎重に読み込む必要がある。一方で、ハイライトの付いた項のみ表示をすれば（逆にハイライトの付いていない項を表示しないで折込む）、表示された項だけ慎重に読めばよいことから、調査担当者の負担は大きく軽減する（読みやすくなる）と考えられる。

(4) 文書の類似順ソート機能：例：「Deskbee」
(人工知能ツール)

1) 「Deskbee」(人工知能ツール)の基本機能説明

まず、人工知能ツールの例として「Deskbee」の基本機能説明をする。

Deskbee へのデータ投入は、例えば発明の名称、要約、請求の範囲などをExcelにて行う。まずMecab⁽⁶⁾にて形態素解析を行い、そこから名詞を取り出し、番

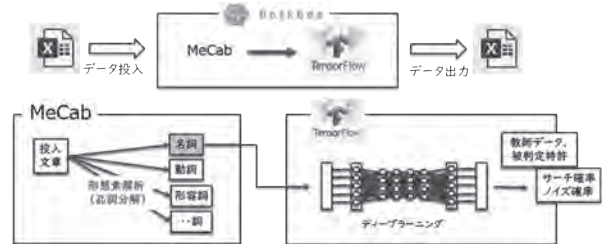


図9：Deskbeeの仕組み1

十分	名詞,形容動詞語幹,****,十分,ジユブン,ジューブン
な	助動詞,***,特殊・タ,体言接続,だ,ナ,ナ
防腐	名詞,一般,***,防腐,ボウフ,ボーフ
性	名詞,接尾,一般,***,性,セイ,セイ
リコピン	名詞,一般,***,リコピン
安定	名詞,形容動詞語幹,****,安定,アンテイ,アンテイ
配合	名詞,サ変接続,****,配合,ハイゴウ,ハイゴウ
可能	名詞,形容動詞語幹,****,可能,カノウ,カノ
な	助動詞,***,特殊・タ,体言接続,だ,ナ,ナ
リコピン	名詞,一般,***,リコピン
含有	名詞,サ変接続,****,含有,ガンユウ,ガンユウ
組成	名詞,サ変接続,****,組成,ソセイ,ソセイ
物	名詞,接尾,一般,***,物,ブツ,ブツ
を	助詞,格助詞,一般,***,を,ヲ,ヲ
提供	名詞,サ変接続,****,提供,テイキョウ,テイキョウ
する	動詞,自立,***,サ変・スル,基本形,する,スル,スル

名詞のみ抽出

十分	名詞,形容動詞語幹,****,十分,ジユブン,ジューブン
防腐	名詞,一般,***,防腐,ボウフ,ボーフ
性	名詞,接尾,一般,***,性,セイ,セイ
リコピン	名詞,一般,***,リコピン
安定	名詞,形容動詞語幹,****,安定,アンテイ,アンテイ
配合	名詞,サ変接続,****,配合,ハイゴウ,ハイゴウ
可能	名詞,形容動詞語幹,****,可能,カノウ,カノ
リコピン	名詞,一般,***,リコピン
含有	名詞,サ変接続,****,含有,ガンユウ,ガンユウ
組成	名詞,サ変接続,****,組成,ソセイ,ソセイ
物	名詞,接尾,一般,***,物,ブツ,ブツ
提供	名詞,サ変接続,****,提供,テイキョウ,テイキョウ

部品番号をつける

十分	0001
防腐	0002
性	0003
リコピン	0004
安定	0005
配合	0006
可能	0007
リコピン	0004
含有	0008
組成	0009
物	0010
提供	0011

※既出の部品には同じ番号が

図10：Deskbeeの仕組み2

号変更したうえで TensorFlow⁽⁷⁾ にかけて、類似度を確率値として Excel 出力する。

Mecab 工程について詳細説明する。たとえば『十分な防錆性を有し、且つリコピンを安定して配合可能なリコピン含有組成物を提供する』を形態素解析し、品詞という部品に分離する。

品詞分解した中から名詞のみを抽出し、そのうえで部品番号が付与される。この時、同じ名詞（ここでは「リコピン」）に同じ部品番号が付与される。ただし、二酸化炭素、炭酸ガスのように同じ意味ではあるが表記が異なるものは異なる部品番号が付与されることに注意が必要である。

また教師特許として、正解としたい特許をサーチ教師、正解としたい特許をノイズ教師と、2つの教師特許を用いて、判定したい特許（被判定特許）一件ずつに対し、サーチ教師、ノイズ教師それぞれに近似しているかを示すサーチ確率、ノイズ確率の2つの確率を TensorFlow にて出力する。通常はサーチ確率からノイズ確率を差し引いた「確率差」の大きい順番にソートし読み込む作業となる。

2) 「Deskbee」の結果表示機能

上述のように Deskbee はサーチ教師特許とノイズ教師特許の2つの教師特許を用い、1件の被判定特許に2つの確率を出力する。Deskbee ではその2つの確率値を用いて、被判定特許の分散状態を見える化する。

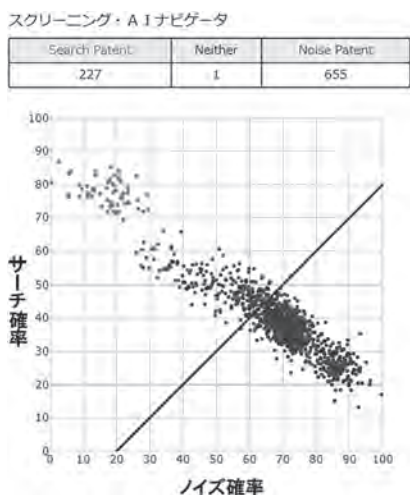


図 11：スクリーン・AI ナビゲータ

図 11 のグラフは「スクリーン・AI ナビゲータ」であるが、横軸をノイズ確率、縦軸をサーチ確率としたもので、グラフ中の「点」は被評価特許と教師特許1件ずつを表す。図中直線の左上のエリアをサーチエリア、右下のエリアをノイズエリアとし、サーチエリアには正解特許、ノイズエリアには正解でない特許が多

く存在するとしている。図 11 には表示していないが、どちらともつかないエリアを Neither エリアとした。

3) 「Deskbee」の初期的課題

図 12 に Deskbee の初期的課題を記す。侵害防止調査 (FTO 調査, クリアランス調査) などの調査実務において、横軸は調査する全特許数を分母とし、スクリーニングした特許数を分子としたときの調査実施率。縦軸は全正解特許を分母とし、スクリーニング中に発見した数を分子とした正解特許発見率。

正解特許が調査する全特許に完全分布している場合は、このグラフでは調査実施率、正解特許発見率で (0, 0) (100, 100) の2点を通る直線となる。

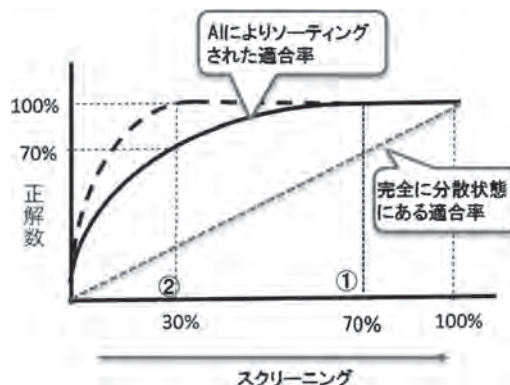


図 12：Deskbee の初期的課題 1

Deskbee により文書の近似順にソーティングした特許の適合率曲線は上に凸の曲線となる。弊社は Deskbee を使って「知財 AI 活用研究会⁽⁸⁾」で一般企業の方々に人工知能ツールを含め効率的知財情報調査の研究⁽⁹⁾をさせていただいている。

初年度の知財 AI 活用研究会の最終報告会において、図 12 のような、ほぼ同じ上に凸の曲線となった事例をいくつかのグループが発表した。しかし、ほぼ同じであるにもかかわらず、2つの相反する意見が出た。

- ① ほしい特許が全体の 70% を読まないで 100% 出てこない。理想は 30% 読んだら 100% 出てくるである。よって Deskbee は使えない。
- ② 30% 読めば、ほしい特許全体の 70% を読むこと

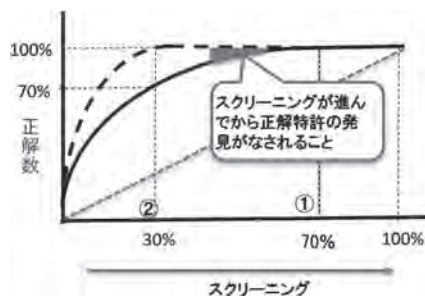


図 13：Deskbee の初期的課題 2

表 2：醜いアヒルの子定理の説明

	体毛色	背の高さ	横幅	体重	目の大きさ	顔の形	足の長さ	毛並み
醜いアヒルの子	灰色	高い	細い	重い	普通	丸	長い	普通
普通のアヒルの子1	黄色	低い	細い	軽い	普通	丸	長い	普通
普通のアヒルの子2	黄色	低い	太い	重い	普通	丸	短い	まだら
普通のアヒルの子3	黄色	低い	太い	重い	小さい	四角	短い	普通

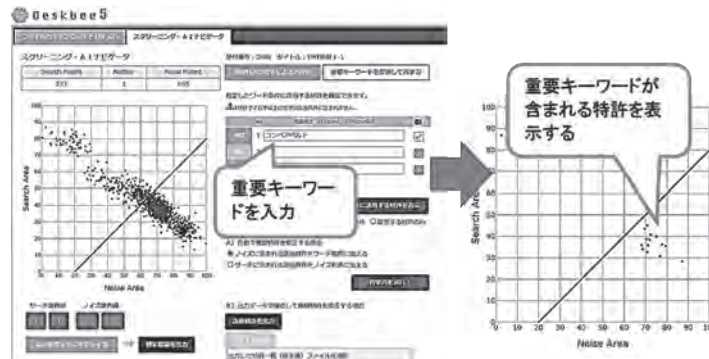


図 14：Deskbee の課題解決

ができる。よって Deskbee は使える。

図 13 のようにスクリーニングが進んでやっと発見される正解を、よりスクリーニングの早い段階で発見できるようにしたい。図 13 に示す楔型部分の正解特許群がなぜ正解なのかを調査したところ、「特許明細書中に正解を決定付ける重要なキーワードが含まれる」ことがわかった。一方で、Deskbee に投入した文書から「醜いアヒルの子定理⁽¹⁰⁾」により、スクリーニングの早い段階で発見は困難であることもわかった。

醜いアヒルの子はなぜ醜いのかについて、物語を知っていれば羽毛の色が「灰色」であるからと答える。この場合、醜いかどうかを決める重要キーワードは「灰色」である。しかし、物語を知らなければ、醜いかどうかの定理は、顔が醜い、背が低い、太っている、毛並みが悪いなど他の文書も含めて総合的に判断することになる。その場合、体毛色よりも他の要素が多く記載されていた場合、そちらを醜い理由とする可能性がある。

特許文書として考えると、特徴を決める項目が多ければ多いほど、つまり使用する特許文書が長ければ長いほど単語数が多くなり、結果他の特長項目に邪魔をされ、決定要素が評価へ占める割合が相対的に低下することが考えられる。

4) 「Deskbee」の課題解決

一般的な人工知能に適用される醜いアヒルの子定理が Deskbee にも当てはまることから、この課題解決のため以下 2 点を機能付加した。

- ① 重要キーワード（構成要件）の検索、抽出機能を付加

- ② 当該重要キーワードを含む案件を教師特許とし、再学習させ「確率差」の下位から重要キーワードを含む特許、及びそれらに近似する案件を除く機能付加

初期的機能としては、スクリーニング・AI ナビゲータは、サーチ確率、ノイズ確率の 2 つの値を利用し、Deskbee 判定の分散状態を確認する機能でとどまっていたが、上記機能付加により、図 14 のように重要キーワードを入力すると、キーワードが含まれる特許だけを表示させることができ、またその表示した特許だけ閲覧することができ、なぜそのような評価値が出たのかの確認をすることもできるようにした。確認作業により、当該特許が重要である場合、教師データとし「再学習」させることができるようにした。

また、重要キーワードのヒントとして、図 15 のようにサーチエリア、Neither エリア、ノイズエリアに位置する特許文書から複合語を取り出し、かつ、他のエリアにない、そのエリア特有の複合語を表示させる機能を付加した。また、全ワードを Excel に出力する機能を付加し、それぞれのエリアにあってほしくない複合語のチェックなどに活用できるようにした。

5) 「Deskbee」の特長

図 16 に混同行列とスクリーニング・AI ナビゲータの関係を示す。スクリーニング・AI ナビゲータの左上のエリアはサーチエリアで Deskbee でサーチと判定された特許が位置するところである。ここには混同行列で TP（真陽性）と FP（偽陽性）が含まれる。一方、スクリーニング・AI ナビゲータの右下のエリアはノイズエリアで Deskbee でノイズと判定された

◎ユニークワード ○全ワード

ワード出力

1 - 300 of 6762		1 - 33 of 33		1 - 300 of 27131	
1 2 3 4 5 >>>		1 2 3 4 5 >>>		1 2 3 4 5 >>>	
住	サーチエリア	住	Neitherエリア	住	ノイズエリア
31	組合せ計量装置	1	ウインチドラム	27	前記ブーム
29	計量ホッパ	1	ウインチ装置	25	シート
21	計量物	1	トルク基準値	25	分野
12	前記計量値	1	トルク制御	24	優先権
11	組合せ秤	1	モータ回転計	23	帯状
10	トラフ	1	ロープ張力	22	ロボット
9	前記計量ホッパ	1	ワイヤーロープ押さえ手段	21	コンベヤベルト
8	ゲート	1	外付け装置	21	動力
8	計量ユニット	1	巻下速度	19	ブラケット
8	直進フィーダ	1	基準トルク	19	プロセッサ
7	計量皿	1	基準変化率	18	ロープ
7	計量手段	1	急速巻下操作	18	干渉
7	計量容器	1	指標値	18	米国
7	軸重	1	前記ウインチドラム	18	両者
7	前記トラフ	1	前記トルク制御	16	M P a
7	前記計量容器	1	前記モータ回転計	16	チェーン

図 15：ユニーク複合語表示機能

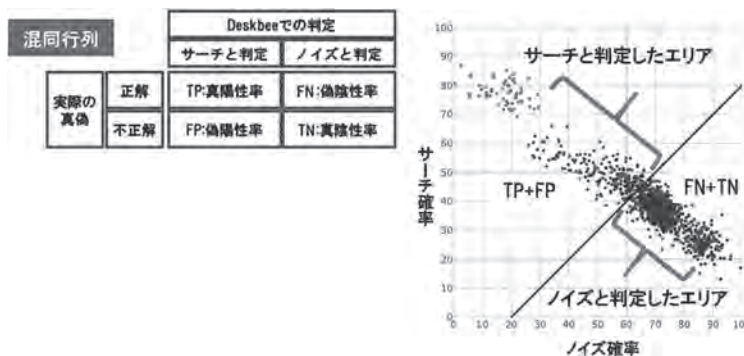


図 16：Deskbee の特長

特許が位置するところである。混同行列では FN（偽陰性）と TN（真陰性）が含まれる。

前章 4) のように、ノイズエリアにおいて重要キーワードが含まれる特許をサーチ教師特許として再学習をかけることにより、FN（偽陰性）が存在しなくなるはずである。

侵害予防調査、SDI 調査など、検索結果として出力された特許はすべて読み込む調査においては、その重要な特許は少量で、ほとんどが不要な特許である。この場合、ノイズエリアから FN（偽陰性）を除いて、読まなくてもよい特許数を多くしたほうが、逆に FP（偽陽性）を除くより、はるかに調査の効率化は向上する。

(5) 『全文日本語翻訳』機能：例「Japio-GPG/FX」

特に研究開発者から「私は外国語の勉強をしているのではない。もっと効率よく外国特許調査ができないのか」と言われることが多くなっている。英語ですら

日本語を読むより時間がかかってしまうのが現状のようである。しかし外国語を翻訳したくても、人手での翻訳は多額の費用がかかったり、性能の良い翻訳ツールがなかったり、1件ずつしかツールでは翻訳できないなどから、一般的な外国特許調査は、非常に手間がかかるものであった。

Japio-GPG/FX は収録範囲が広く、また日本語訳文が非常に高品質であり、かつ多数の特許を一度にダウンロードすることができるので、外国語の特許を日本特許と同じ日本語で読むことができる。また、日本語で言語を統一することで、以上述べてきた知財リーガルテックを日本語、外国語の区別なく一つの集合とすることで大幅な知財調査業務の効率化を図ることができる。

3. 3 知財調査効率化のための知財リーガルテックについて

表 3 に知財調査効率化のための調査ニーズと、その

表3：調査ニーズと知財リーガルテック

調査ニーズ	適用調査名	ハイライト (バー)	ハイライト ソート	ハイライト 折込み	Deskbee
読みたいものだけを読む	無効調査	△	○	○	△
読みたいものから読む (優先順位付け)	侵害予防調査・SDI調査	×	○	△	○
読みたくないものを除く	侵害予防調査・SDI調査	×	△	×	○
見落としをなくす	無効調査	○	○	○	□

○：該当、×：該当しない、△：一部該当、
□：調査範囲を拡大することができることから調査範囲外にある重要特許を発見可能

ニーズを満足するための知財リーガルテックをまとめた。重要キーワードを用いたハイライト系は、無効調査などに有効、Deskbeeは侵害予防調査、SDI調査向き^(*)であることがわかる。

特に人工知能ツールについては高額であることもあり、すべての調査を一つのツールで済ませたいとされる方が多いが、ツール適性に応じた調査手法（餅は餅屋）をとるべきと思われる。

(*) 日本化薬株式会社知的財産部の奥村公人がDeskbeeを利用したSDI調査の有用性について知財AI活用研究会で成果⁽⁹⁾発表した。それによると過去に蓄積した「SDIで評価した特許」をサーチ教師特許、ノイズ教師特許として自己評価させることで、サーチ教師特許、ノイズ教師特許の精緻化がなされ、それを教師として使用するとその後のSDI調査において発見すべき特許がサーチエリアに抽出することができたとのこと。DeskbeeがSDI調査向きである一つの検証結果である。

4. まとめ

企業内弁理士だけではなく、弁理士事務所の弁理士に、研究開発そのものの質・量ともに高め、結果的に発明の質・量が向上するために、研究開発部門と研究開発工程においてできるだけ早い段階でのコミュニケーションを今よりも増していただけるよう期待を書かせていただいた。

またそのコミュニケーションには、弁理士としての従来の知識・スキルだけではなく、知財リーガルテックの知識を持ち利用（支援・提案）できることが大切である。そしてそのためには研究開発部門におけるニーズの把握、それにあった有用な知財リーガルテックの選定眼を持つことが大切である。

今回は研究開発力強化と知財調査効率化のための知財リーガルテックを紹介した。本著作成にあたり、インタビューで多くの示唆を与えていただいた弁理士の皆様に感謝する。

(参考文献)

- (1) 特許庁ホームページ 特許行政年次報告書 2020年版〈統計・資料編〉 https://www.jpo.go.jp/resources/report/nenji/2020/index.html#toukei_shiryou
- (2) 鶴見隆, Japio YEAR BOOK 2010, 25th ANNIVERSARY pp72～75
- (3) 中村栄, 情報管理 2008 vol51, No.7, pp457～468
- (4) 特許庁, 戦略的な知的財産管理に向けて～技術経営力を高めるために～〈知財戦略事例集〉 2007年4月, pp147～159
- (5) UMethodのご提案 <http://ipfine.com/Download/UMethod.pdf>
- (6) Mecab (和布蕪) <http://taku910.github.io/mecab/>
- (7) TensorFlow <https://www.tensorflow.org/?hl=ja>
- (8) 知財AI活用研究会 <http://www.ipfine.com/deskbee/AI.html>
- (9) 平尾啓, 知財AI活用研究会の研究事例紹介, 情報の科学と技術 2020年70巻7号 pp349-354
- (10) 安藤俊幸, 機械学習を用いた効率的な特許調査方法, Japio YEAR BOOK 2018, pp238～249

(原稿受領 2020.11.30)