

人工知能が知財業務に及ぼす影響

国立情報学研究所 教授・理学博士
 ランドン IP 合同会社 シニアディレクター
 日本アイ・ビー・エム株式会社 主席研究員・工学博士
 会員

宇野 毅明
 野崎 篤志
 那須川 哲哉
 小川 延浩

要約

人工知能が話題となっている昨今、知財業界にもその影響は波及しつつある。特許庁では、「人工知能技術を活用した特許行政事務の高度化・効率化実証的研究事業」の企画提案の公募を行い、平成 28 年度から人工知能技術の活用検討を始めている。^(*1) FRONTEO(旧 UBIC)社からは人工知能を使った特許情報調査・分析システムがリリースされ、話題となった。^(*2) また、野村総研からは、弁理士という職業が機械にとってかわられる可能性が高いという研究報告が公表されている。^(*3) このように、人工知能関連の技術が世の中に浸透していく中で、知財業務は今後どのように変化をしていくのだろうか。本稿では、人工知能が今後、知財業務にどのような影響を及ぼすかについて考察する。

(*1) https://www.jpo.go.jp/koubo/koubo/jinkou_chinou.htm

(*2) <http://www.kibit-platform.com/products/patent-explorer/>

(*3) http://pc.watch.impress.co.jp/docs/news/20160112_738555.html?ref=rss

目次

1. はじめに
2. 人工知能概論
 - (1) 人工知能とは何か
 - (2) 人工知能研究の歴史
 - (3) 直近の研究動向
 - (4) 人工知能技術の今後
3. 人工知能が知財業務に及ぼす影響
 - (1) 雇用の未来
 - (2) 機械との競争, 機械との共存
 - (3) 知財業務についての機械と人間の棲み分け
4. まとめ: 知財業務に携わる者は今後人工知能とどう関わっていくのか

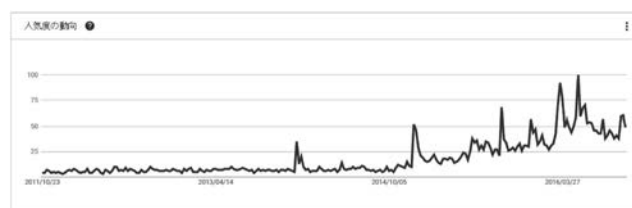


図1 Google Trends で「人工知能」を調べた結果

1. はじめに

ここ数年は人工知能ブームであり、ニュースや雑誌の特集、新刊書籍の出版など、多くのメディアに人工知能がテーマとして取り上げられている。例えば Google Trends で人工知能というキーワードを調べてみると、2014 年以降、大きな伸びを見せてきており(図1)、この勢いはまだしばらく続くと考えられる。

このような背景の中、日本弁理士会・関東支部では、「人工知能時代の知財業務について考える ～特許検索・分析業務をケースとして～」と題して、2016 年 7 月 8 日に会員向けの研修会を実施した。本稿は、この研修会で行った講演と議論の内容をベースに人工知能と知財業務の関係をまとめあげたものである。まず、2 章では、議論の前提知識としての人工知能概論について説明する。人工知能とは何か、どのような歴史を経て技術が発展してきたか等の一般的な説明を行う。次に 3 章では、人工知能関連技術の発展によって、知財業務がどのような影響を受けるかについて考察する。人工知能によって雇用が奪われる等の議論があるが、知財業務についてはどうかについて、特許情報検

索・分析業務をケースに考察する。そして最後に4章で本稿のまとめを行う。

2. 人工知能概論

(1) 人工知能とは何か

人工知能学会のウェブサイトによると、「人工知能の概念自体は、1947年の「Lecture to London Mathematical Society (ロンドン数学学会での講義)」にてアラン・チューリングによって提唱されたとするのが良い」と説明されている⁽¹⁾。しかし、「人工知能」という言葉の定義は専門家の間でも定まっておらず、例えば人工知能学会誌において、13人の専門家が説明するとそれぞれ異なる定義となることが示されている⁽²⁾。とはいえ、究極的には、人間の知能と同じモノをコンピュータ、つまり算術演算と数値の記憶の組合せだけで実現することを目指していると考えられる。言葉の定義は不明確であるものの、アラン・チューリングによって、「チューリングテスト」と呼ばれる「人工知能であるかどうか」を判定する方法が提唱されており、2014年6月に、はじめてロシアのスーパーコンピュータが、この判定に合格したとして話題となった⁽³⁾。この「チューリングテスト」は、「コンピュータと判定者である人間が、例えばキーボードとディスプレイ等を介して会話をを行い、判定者が相手をコンピュータと見抜けなかったら合格」というものである。人工知能をどのように実現しているかという中身には一切触れず、会話という外側からだけで判定するところにチューリングテストの特徴がある。人工知能を実現するためには、必ずしも人間の脳などを模倣する必要はなく、外から見て人間と同じように振舞うことができれば良いということになる⁽⁴⁾。

(2) 人工知能研究の歴史

人工知能研究は、これまでブームと冬の時代を繰り返してきており、今回で3度目のブームと言われている⁽⁵⁾ (図2参照)。

第1次AIブームは1956～1960年代であり、この時期は「探索・推論の時代」と言われている。この当時には、ニューラルネットワークや自律移動ロボット、自然言語対話システムなど、人工知能技術の基礎が作られた。しかし、探索と推論を中心としたシステムでは特定の単純化された問題を解けるに留まり、現実存在する複雑な問題との間にはまだ大きな隔たりが

あった。このように当時の技術の限界が明らかになるにつれてブームはおさまり、冬の時代を迎えることになる。

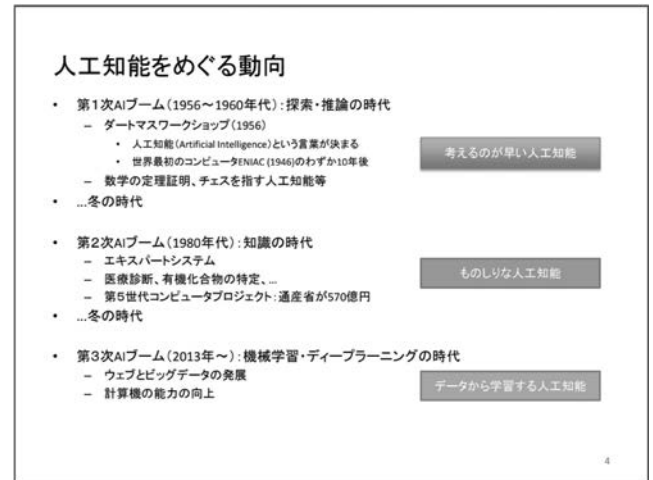


図2 人工知能をめぐる動向⁽⁶⁾

その後、1980年代に次の第2次AIブームである「知識の時代」がきたと言われている。この時期には、専門家の知識をコンピュータに入力しておき、その知識をベースに推論を行わせ、意思決定の支援を行う「エキスパート・システム」の研究が発展した。しかし、知識を記述・管理することは困難かつ膨大な労力を払うため、やがて行き詰まりを見せてブームはおさまった。再び冬の時代へと突入する。

そして、2013年以降の第3次AIブームへとつながり、これが現在も続いている。今回のブームは、ビッグデータの時代に広がった機械学習と、技術的に大きなブレイクスルーとなったディープラーニングの2つの大波が重なって生まれている。第2次AIブームと比較した第3次AIブームの特徴は、知識を学習する能力にある。第2次AIブームでは、多様な知識を与えて推論させることで知的な能力を実現できるようになり、人工衛星の制御など、従来は一部の専門家が知識を駆使して行っていた作業にAIが使われるようになったのが、エキスパート・システムと呼ばれた所以である。しかし、当時の知識は、システム向けに定義された特殊な形式で記述されており、一部の専門家が慎重に作成する必要がある、作成の労力や、記述能力の柔軟性などの問題が行き詰まりの原因となった。それに対し、第3次AIブームでは人間が通常扱っているデータから知識を学習している。例えば、『ジェパディ!』(Jeopardy!)という1964年から続いている米国の人気クイズ番組で、74連勝という最多連勝記録保持

者のケン・ジェニングス氏と、300万ドルを超える累積最高賞金記録保持者のブラッド・ラッター氏という、番組視聴者にとっては非常に有名な二人と、IBMのワトソンというコンピュータのシステムが2011年に対戦し、勝利した。米国のテレビ番組のため、日本においては、一般の人々の間で、さほど話題にならなかったが、米国においては、博学で頭が良いというイメージの強いクイズ王に対するコンピュータの勝利が、視聴者に強いインパクトを与えた。このインパクトが、第3次AIブームの一つのきっかけとなっている。

IBMのワトソンに勝利をもたらしたのは、人間が作成したルールではない。長寿番組のため、過去のクイズ問題と正解のデータが大量に存在するものの、同じ問題は二度と出ないため、過去の問題と正解を与えても新しいクイズには解答できない。勝利したのは、ワトソンが、Wikipediaや、ニュース記事・ブログなどのほか、聖書や歌の歌詞といった大量のテキストから知識を学習することができたためである。近年、精度が劇的に向上して第3次AIブームのベースとなっている画像認識や音声認識のベースになっているのも大量のデータである。例えば、ネット上に大量に存在するキャプション付きの写真が、高度な画像認識技術を可能にしている。また、自動運転の実現に重要な役割を果たすのが膨大な走行データであり、その大部分は、一般人が運転している自動車から収集するデータである。すなわち、コンピュータのために専門家が慎重に作成したルールではなく、人間のために用意された膨大なデータや、一般の人々が日々の活動で作りに出しているデータを活用できるようになったのが、第3次AIブームの技術の特長であり、いわゆるビッグデータがその原動力となっている。

第2次AIブームのエキスパート・システムには、専門家がルールを保守しなければ使い物にならないという問題があった。それを考えると、第3次AIブームの技術がより大きな可能性を秘めていることは想像に難くないのではないだろうか。

(3) 直近の研究動向

①データ収集環境の改善

前述の通り、人工知能関連の研究は、膨大なデータをいかにして活用するかが中心となってきた。人間が通常使っている言葉で記述されたテキストデータ

を処理する自然言語処理の分野においては、第2次AIブームが下火になった1990年代から、大量のデータを統計的に処理する取組みが次第に盛んになってきた。1980年代までは一般家庭やオフィスへのPCの普及が進んでおらず、電子化されてコンピュータで扱えるテキストデータの量自体が少なかった。コンピュータの計算能力も現在と比べると桁違いに低い状況であった。そのため、現在主流となっている統計的機械翻訳手法も、当時はまともな翻訳結果を出すことができず、1990年代前半までは学会でも懐疑的な見方が多かった。しかし、インターネットとPCの普及と共に、電子化されたテキストデータが大量に使えるようになり、高度な専門的能力を駆使して作成した知識に基づく手法よりも、データから機械的に学習した知識に基づく手法が優勢になっていった。統計的な機械学習法に基づく処理を実現するためには、基本的には、タスクに応じた入力と出力のデータが必要となる。例えば、英日機械翻訳であれば、入力となる大量の英文と、各英文に対する日本語文を集めた大量の対訳データが必要である。また、文中の単語の品詞を判断したり、「長野」が人名を示しているか地名を示しているかといった言葉の意味の曖昧性を解消したりする処理を実現するためには、文の単語に品詞や意味情報を示すラベルを付与した大量のデータが必要となる。こういったデータの作成も従来は専門家の手に委ねられていたが、インターネットやPCが普及した結果、ネット上の不特定多数の人々にその作業を依頼するクラウドソーシング(crowd sourcing)で作成できるようになり、データ作成のコストを下げることによって、データ量を増やしやすくなった。

②膨大なデータ処理に対する効率化向上

2000年代に入ってから、データ量の増加とコンピュータの計算能力の向上を背景に、機械学習の研究が盛んになり、サポートベクターマシン(SVM)などの技法が開発されて普及し、自然言語処理の精度を上げると共に、テキストマイニングや評判分析・質問応答など、その応用を広げてきた。当初の機械学習では、こういった特徴量を用いて学習させるかが重要であり、例えば、日本語の品詞を判断させるためには、前後の単語の文字種や句読点を考慮するなど、様々な特徴量の有効性を検討することが、精度向上の鍵となっていた。そのため、機械学習では試行錯誤が重要

であり、大量のデータを処理する膨大な計算を現実的な時間内に完了させることが必要となる。そのためには、処理を並列化させることが必要となる。そこで注目されたのが GPU (Graphics Processing Unit) である。この元々はコンピュータグラフィックスの描画を高速化するチップであった GPU を活用することで機械学習が効率良く行えるようになってきた。

③新たな手法ディープラーニングの出現

スマホなどの新しいデバイスの普及を背景に CGM (Consumer Generated Media) などネット上の情報が益々拡大し、テキストのみならず、画像や動画、音声など、アクセス可能なデータは増える一方である。こういった状況を背景に、2010年代に入って脚光を浴びるようになったのがディープラーニング (深層学習) である。第二次 AI ブームでも着目されていたニューラルネットワークを多層化して、大量のデータで学習させることで、画像認識や音声認識では他の手法を大幅に上回る精度を実現し、人の認識や音声のテキスト化などを実用レベルに引き上げることができるようになった。また、従来の機械学習では、用いるべき特徴を専門家が検討していたのに対し、ディープラーニングでは、特徴を自動的に判断するところに大きな進化がある。データをどんどん投入しさえすれば精度の高い処理が自動的に実現される仕組みであり、それが、ディープラーニングへの期待が高まる理由でもある。但し、だからといって、専門家が不要になるわけではない。ディープラーニングを構築する上で、ネットワーク構成や層の深さを調整する必要があり、こういったタスクを設定し、どのようなデータをどう投入するかなど、多くの課題を検討する必要がある。ディープラーニングに関する高いスキルを持った人材が求められる。画像認識や音声認識で従来手法の精度を大きく上回ったディープラーニングであるが、テキストを対象とする自然言語処理においては、ディープラーニングが既存手法の精度を大きく上回るという結果を出せていないのが、現時点での状況である。膨大な量の画素から構成される画像や、音波の波形で表わされる音声と異なり、テキストは、それ自体が抽象化された文字列の集まりであるため、今まで誰も気付いていなかった特徴量がディープラーニングによって活かされるという状況になり難い可能性がある。それでも、ここ 1~2 年の間に研究の最前線では、ディープ

ラーニングの取組みが急激に増え、次第に成果を出しつつあるという状況である。

④自然言語処理技術の発展

現在の自然言語処理では、実世界と結びついた言葉の意味を扱っているわけではなく、あくまでもテキストにおける言葉の使われ方を分析し利用しているに過ぎない。例えば、「りんご」という言葉を扱う上で用いる情報は、実際にりんごを噛んだときの感触や味覚といった人間の体験と結びついた情報とは全く異なる。大量のテキストデータにおいて「りんご」という言葉が、他のどのような言葉とどのような関係で使われているか、例えば、「食べる」や「買う」の目的語になったり、「赤い」や「甘い」「すっぱい」という形容詞で修飾されたりするといった情報である。このような情報であっても、膨大なデータを対象とすると、言葉の意味を扱っていると捉えられるような処理ができるようになってきている。例えば、近年注目されている Word2Vec は、ある言葉が、どのような言葉と共起しているかという情報を膨大なテキストから抽出し、ベクトル化して扱う手法である。言葉の種類だけの次元を用意するのは計算量的に現実的でないため、次元圧縮している。単純な情報をベースにしているが、意味が似ている言葉を集める類似度計算や、「男性」と「彼」の関係を「女性」に適用して「彼女」を導き出すといった意味演算への期待から、活用が広がってきている。

(4) 人工知能技術の今後

これまで起こった 2 度の人工知能ブームはいずれも何らかの壁にあたって収束をむかえてきており、今回の 3 度目のブームも同じ道をたどるのではないかとの疑念を抱く人も多いだろう。しかし、ブームが仮に終わったとしても、人工知能の研究は終わらず、引き続き発展することは間違いない。コンピュータの性能向上により、2030 年代初頭には 1000 ドルで買えるコンピュータの処理能力はおおよそ人間の知性と並ぶといわれている⁽⁷⁾。そして、2045 年にはコンピュータの処理能力は人間の知性を大幅に超え、破壊的な変化が生じる (技術的特異点) と言われている。ブームが終わるか否かに関わらず、今後、コンピュータ化や人工知能の導入によって、産業構造やワークスタイルが変化することは避けられない。

3. 人工知能が知財業務に及ぼす影響

(1) 雇用の未来

2013年、「コンピュータによって影響を受けやすい職業は何か」をテーマとした論文「雇用の未来⁽⁸⁾」が、オズボーン氏らによって発表された。本論文では、702の職業について、ロボット等の機械に取って代わられる可能性を数値で示した点が大きな特徴と言える。そして、その考察によると、米国雇用のうち47%が将来的にリスクにさらされると述べられており、その刺激的な内容からも話題となった。特に、論文中に「弁理士」(本文中では“patent lawyers”)という記載があったため、知財業界では興味を持たれた方も多かったのではないだろうか⁽⁹⁾。この論文の中では、今後コンピュータ化が法律やファイナンスの分野で進み、弁理士等が現在行っている業務がだんだん高度なアルゴリズムに取って代わられるだろうと説明されている。ただし、具体例として挙げられているのは、シマンテック社のClearwellシステムが、eDiscoveryの情報開示プロセス管理において、膨大な文書の分析やソーティングを高速に行えるようになったという事例である。明細書作成や中間処理といった弁理士の中核業務については触れられておらず、弁理士という業務が機械に取って代わられるかについては更なる考察が必要と考えられる。

(2) 機械との競争, 機械との共存

このように、技術革新によって仕事や雇用が失われるということは、今回のコンピュータや人工知能の登場によってはじめて議論されるようになったわけではない。電話交換手、植字工などのように、これまでの技術革新によって自動化等が進み、なくなってしまった職業も存在する。しかし、今回のコンピュータ・人工知能による労働の代替は、これまでのように肉体労働や単純労働を代替してきた工学的な技術革新とは異なるタイプであると、国立情報学研究所の新井教授は指摘する⁽¹⁰⁾。同氏によると、20世紀後半に起こった情報科学という技術革新により、21世紀において、その職を追われる可能性があるのは、知的労働を行う、いわゆるホワイトカラーということになる。それでは、ホワイトカラーが担う知的労働すべてが機械に置き換わるかというところではない。労働において新技術が活用される場合、文字通り人間を機械に置き換えるのではなく、組織やプロセスを見直して再編

し、再設計するというのが最善の方法である⁽¹¹⁾。その際、新技術によって不要となってしまう労働は、多くの場合は定型業務である。米国では、ターボタックスという確定申告作成ソフトが普及したことにより、確定申告作成サービスを提供してきた大勢の人が駆逐された。電子申告の導入などでワークフローが変化すると共に、人工知能をはじめとした高度なアルゴリズムが追加されることで仕分けなどの従来は人間が判断していた部分も自動化できるようになり、確定申告作成サービスの大部分が定型業務となってしまったのである。この例に限らず、例えば空港などの案内業務、経理事務、特定の業務のみを受け付ける窓口業務などは同じくここで言う定型業務に該当すると考えられる。一方で、定型業務に落ちない部分には人間にもまだチャンスがある。特に、新しいアイデアを思いつくことなどの発想力を要する業務はコンピュータには困難であるため、人間のほうが比較優位となる。コンピュータは枠にはめられたパターン認識には優れているが、少しでも枠をはみ出すとまったくお手上げになるからである⁽¹²⁾。アイデアを作り出すことを仕事とする研究者などに限らず、アパレルの接客業や、居酒屋の店員のように一見定型業務を行っているように見えても、既知のルール以外の対応を臨機応変に行う必要のある業務は直接的にはアイデアを作る仕事でなかったとしても同じくこの類型に含まれると考えられる。このように双方に得意不得意があることから、コンピュータが苦手とする部分を人間が担い、コンピュータとペアを組むという考え方がある。チェスの絶対王者だったカスパロフが1997年にIBMのスーパーコンピュータ、ディープブルーに屈したことは有名であるが、2005年に当時最強のチェス・ソフトウェアHydraは、強いプレイヤーとありきたりのノートパソコンのチームに負けてしまう。そして、決勝戦で最後に勝利したのは、グランドマスターと強力なコンピュータのチームのほうではなく、二人のアマチュアプレイヤーと三台のPCのチームであった。二人は対戦相手の裏をかく手を積極的に打つことで勝利した。弱い人間であってもコンピュータを用いて強いプロセスと協力すれば、強い人間が弱いプロセスと協力するよりも上回ることがあることを示した⁽¹³⁾。

(3) 知財業務についての機械と人間の棲み分け

本節では知財業務における人工知能と人間の棲み分

けについて述べる。知財業務といっても出願・権利化、調査・分析、渉外や管理業務など多岐にわたる。まずは2016年3月に発表された「人工知能技術を活用した特許行政事務の高度化・効率化実証的研究事業」⁽¹⁴⁾から特許庁が人工知能を利用してどのような業務の効率化を図ろうとしているか確認する。

本事業は「特許行政事務への人工知能技術の適用可能性の検討 出願等の受付、方式審査、分類付与、実体審査のそれぞれの業務プロセスについて、現在の業務内容を下記の担当課へのヒアリング等を行って把握し、当該業務内容について人工知能技術の適用可能性を検討する」ことを目的としている。ここで対象となっている業務は下記となっている。

- 出願等の受付：出願課
- 方式審査：審査業務課
- 分類付与：調整課（特許，実用新案），意匠課（意匠），商標課（商標）
- 実体審査：調整課（特許，実用新案），意匠課（意匠），商標課（商標）

前述の通り、第3次AIブームは機械学習とディープラーニングの進展、そして学習対象としての膨大なデジタルデータの蓄積に支えられている。特許情報においてもCD-ROM公報が発行された平成5年以降だけでも約840万件の公開特許公報が発行されている⁽¹⁵⁾。過去の膨大な蓄積からルールを抽出することにより、出願等の受付や方式審査などの定型業務についてはAIで効率化できる余地が大きいと考えられる。

また特許・実用新案へのIPCやFI・Fターム分類の付与についても、既に公開されている公報テキストデータ（主に【要約】および【特許請求の範囲】）と付与されている分類データから一定のルールを導くことは不可能ではない。このルールを特許庁が受理した出願明細書へ適用することで分類付与作業の効率化および分類付与のさらなる高精度化を図ることができると予想される⁽¹⁶⁾。また、米国特許商標庁および欧州特許庁主導で2013年1月から利用開始したCPC（欧米共同特許分類）に対して日本特許庁は正式な採用を表明していないが、このCPCの日本特許・実用新案公報への付与もAIを活用すれば付与コストや付与に必要な工数面で大きなメリットがあると考えられる。

最後の実体審査では主に先行技術調査および拒絶理由通知・拒絶査定判断などへの人工知能援用を検討するものと思われる。以降、知財業務の中でも特に先行技術調査（特許調査・分析業務）への人工知能の適用および機械と人間との棲み分けについて考察していく。

特許調査および分析業務において人工知能が着目されたのは、2015年11月にFRONTEO（旧UBIC）が「Lit i View PATENT EXPLORER」の発表がきっかけである。このシステムは、発明提案書や無効化した特許などの見つけたい文書の内容を教師データとして、まず人工知能に学ばせる。その教師データを基に人工知能が、対象の母集団を解析し、スコアリングして教師データとの関連性が高い文書がスコア順に並ぶものである（図3）。

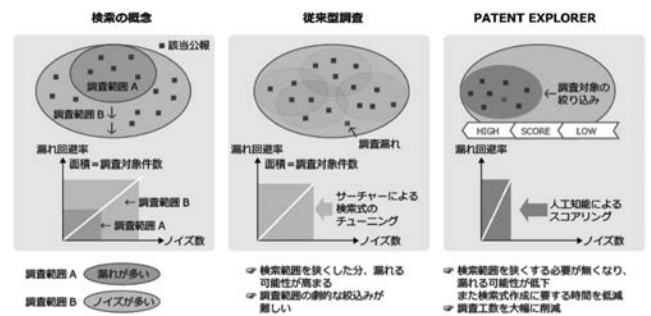


図3 検索の概念およびFRONTEO社 PATENT EXPLORERの特徴⁽¹⁷⁾

人工知能といっても、見つけたい文章に対してズバリ回答するものではなく、関連性が高いものをスコア順にソートする機能である。しかし、人工知能を援用することによって特許文献のレビュー効率は確実に向上すると言える。また、図4に示すように従来型調査では、調査漏れを防ぎつつ該当公報を抽出するために、サーチャーが検索式作成を工夫するのが常であったのを省略・省力化するものである。

上述のように検索式という母集団を形成した上で、目的に対応して特許調査・分析を実施する。図3に特許調査・分析の目的・種類を示した。特許調査・分析といっても知財サイクル（創造，保護，活用）およびR&Dサイクル（研究，開発・設計，生産・事業化，販売・マーケティング）に合わせて様々な種類がある。

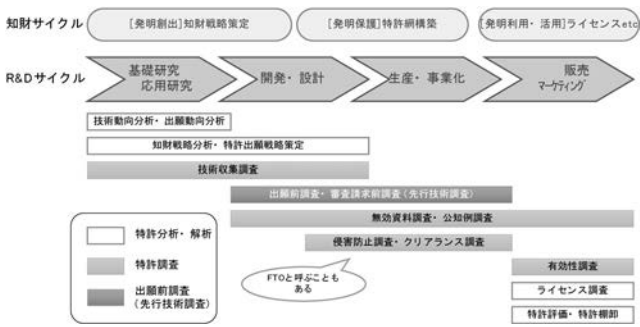


図4 企業の知財・R&D活動と特許調査・分析の種類

よりマクロな視点で特定技術について動向や特定業界等における競合他社の出願戦略を把握するために実施する技術動向分析・競合他社分析，研究前段階・研究中に他社の先行特許について収集する技術収集調査，他社登録特許の無効化，公開特許の権利化阻止をするための無効資料調査・公知例調査，自社製品を上市する際に他社登録特許を侵害していないか確認するための侵害防止調査（またはクリアランス調査，FTOと呼ぶこともある），そして出願人サイドで特許出願前・審査請求前に行い出願の要否・権利化の成否を見極めるための出願前・審査請求前調査（先行技術調査）である。

上述のように，特許庁では実体審査における先行技術調査への人工知能の適用可能性についても現在検討を進めている。様々な種類の特許調査・分析が存在するが，図5に示すように業務フローとしては共通する部分も多い。

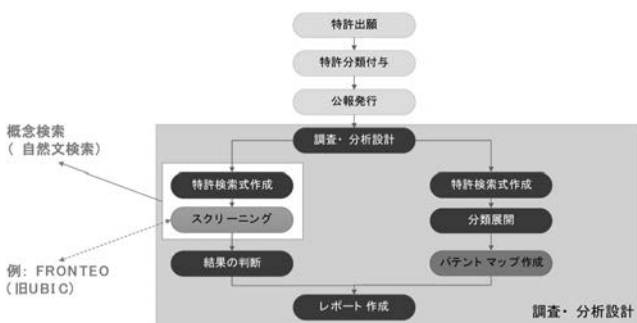


図5 特許調査・分析の業務フロー

例えばどのような調査・分析を行うべきかという設計フェーズ，その調査・分析設計を基に母集団を確定する検索式作成フェーズ，検索式でヒットした公報を読み込むスクリーニングフェーズ，ノイズ除去後の公報に対して分類軸（無効資料調査であればクレームの構成要件，動向分析であれば課題，解決手段）を付与

する分類展開フェーズ，動向分析であれば分類展開結果を可視化するパテントマップ作成フェーズ，そして最終的に結果を取り纏めるレポート作成フェーズである。

図6には著者の私見として，各フェーズにおけるタスクが今後も人間が強みを発揮する領域なのか，人工知能つまり機械によって置換されてしまう領域なのか，または人間と機械が協調して相互補完しあうことが好ましい領域なのか，の3つのいずれかをまとめている。

調査・分析タスク	人間	AI	人間・AI協調
調査・分析設計	○		
特許検索式作成			○
スクリーニング (母集団のノイズ・非ノイズチェック)			○
結果の判断 (新規性・進歩性・侵害・非侵害など)	○		
分類展開 (課題・目的・技術項目等への展開)			○
定型化されたパテントマップ作成 (スコアリングなども含む)		○	
非定型のパテントマップ作成	○		
レポート作成			○

図6 特許調査・分析業務における人間と機械の棲み分け

まず第3次人工知能ブームを牽引している機械学習やディープラーニングが発達しても，人間が引き続き強みを発揮する領域としては「調査・分析設計」と「結果の判断（新規性・進歩性，侵害・非侵害）」の2つが挙げられる。人工知能が効果を発揮するためには，過去の膨大な蓄積から一定のルールを導くことが必要であるが，ゼロベースで目的から仕様までを定めたり，その都度変化する内部環境・外部環境に応じた最適解を導出することは困難であると推察される。そのため，この2つの領域については「人間が強みを発揮」と判断した。

次に現在人間が行っている業務であるが，ゆくゆくは人工知能に置換されてしまう領域としては「(定型化された)パテントマップ作成」が挙げられる。パテントマップとは特許情報を調査・整理・分析して可視化したものであり，件数推移やランキングのような統計解析マップもあれば，技術変遷図や構成部位マップのように非統計解析マップもある。機械によって置換されるのは主に統計解析の手法で作成される定型化されたパテントマップである。

最後に人間と人工知能が協調する領域であるが「特許検索式作成」，「スクリーニング」，「課題・解決手段

等への分類展開」,「レポート作成」など多岐にわたる。「スクリーニング」や「課題・解決手段等への分類展開」については上述した FRONTEO 社のシステムで既に機能が実現していると言えよう。今後はより少ない教師データでの精度の高いスコアリングや企業ごとの特徴や過去実施済み案件結果をベースとした企業特有ルールの設定などの高機能化が図られるものと予想される。人工知能の援用によって効率化・高精度化は期待できるが、最終的な各公報の関連性の有無や分類軸との適合性判断は人間が実施すべき領域であると言える。

「特許検索式作成」についても「スクリーニング」と同様で、見つけたい技術的特徴を記載した先行文献やテキストを教師データとして入力することによって、母集団を形成し、例えばその母集団の中でスコアが上位 20% が包含されるようにキーワードや特許分類などの検索キーを組み合わせた検索式案を提示することは技術的に可能であると容易に推察できる。一方で先行技術調査など一部用途に限定されるが、人工知能による「スクリーニング」能力が高精度化およびデータ処理能力が向上されることで、そもそも検索式から生成される母集団という一定の枠を設定せずに、過去発行された全文献を対象に調査・分析を実施することが定常化することも予想される。

本節では特許調査・分析業務を例として人間と機械の棲み分けについて述べてきたが、AI による特許明細書作成・特許翻訳についても、上述の特許調査・分析業務とのアナロジーで AI により実現できるタスクもあるが、すべて機械で置き換えることができるわけではない。しかし、過去の膨大なテキストデータの蓄積をもとに明細書作成や翻訳が効率化・省力化されることは疑いの余地がない。また、特許明細書の対象となる発明そのものを AI で創出することが可能か否かという点については本稿の対象範囲外となるため、省略させていただく点ご了承いただきたい。

4. まとめ：知財業務に携わる者は今後人工知能とどう関わっていくのか

本稿では、機械が人間に勝る部分をうまく活用し、人間が更に生産性を上げられるように協調するにはどうすべきかについて考察してきた。このような理想がある一方で、これまでに機械によって代替された結果、無くなってしまった職業があることから目をそ

むけてはいけない。最近の人工知能ブームで、必要以上に危機感を煽るような記事がみられることも確かではあるが、社会構造上、新技術の影響を受けて変化する部分としない部分があることを冷静に見ていく必要がある。また、人間と機械が協調するという場合においても、人間が機械を活用するという人間主体の協調もあれば、機械中心のワークフローの一部に人間が入るといった機械主体の協調もある。人間と機械が協調する世界が、必ずしも人間にとって都合の良い未来であるとは限らず、その点についても冷静に状況を見ていく必要がある。

今、人工知能技術の周辺分野として、「ヒューマンコンピューテーション⁽¹⁸⁾」という学問分野があり、この分野は先の協調の例で言うと後者に近い。多少語弊があることを認めつつも説明を加えると、この分野の研究はコンピュータがタスクを行うために、コンピュータにとって苦手な部分を人間に補完させるアプローチとみることもできる。その一例として有名なのが、「reCAPTCHA」というシステムである。ウェブサービスにおいて、利用者が人間であることを確かめることで、コンピュータプログラムによる過剰なアクセスなどを防止する「CAPTCHA」という認証システムがある。コンピュータには判別しにくい捻じ曲がったような数字や文字が現れて、その入力を求め、結果を照合することで認証を行うシステムである。「reCAPTCHA」は、この「CAPTCHA」をうまく応用し、OCR（自動文字認識）プログラムで認識できない場合を補うというシステムである。このシステムによって、コンピュータが不得意なケースが補えるだけでなく、コンピュータが認識できないケースに対する正解データを収集し、更にコンピュータの認識精度も向上させることができる。当初はコンピュータにとって苦手とされていた処理が、こういったシステムの活用によって、コンピュータの能力が人間に追いつく可能性がある。

ここでもう一度、本節の主題である人間と人工知能の関わり方に戻す。本稿では、ここまでで、知財業務において、人間と人工知能が今後どのように協調していくか、分担していくかについて検討を深めてきた。あるタスクは現時点では人間でない十分な質が担保できないことや、機械の方が向いていることなどが主観的ではあるものの考察した。しかし、人間の方が向いているというだけで必ずしも人間の業務として残る

かはわからない。その業務のやり方や質の評価が、人工知能で取り組みやすいものになってしまったり(ソフトのマニュアルなどにある、精度の悪い翻訳)、その業務自体が、人工知能によって簡単にできてしまう他の業務でおきかわってしまったりすることがあるためである。今後、我々は知財業務に関わる者として、人間と機械とがどのように協調していきたいか、という全体システムの設計をしていく必要がある。

※免責

本稿に記載されている内容は、全て、著者個人の見解に基づいており、著者の所属組織の見解を示すものではありません。

以上

(注)

- (1)人工知能学会ウェブサイト (<https://www.ai-gak-kai.or.jp/whatsai/AItopics5.html>)
- (2)松尾豊,「人工知能は人間を超えるか」, 角川 Epub 選書, p45
- (3)2016/6/8 にレディング大学で開催された「Turing Test 2014」においてコンピュータ「Eugene」がチューリングテストに合格したと話題になったが、専門家の間ではこの合格判定について反論もある。
- (4)松尾豊,「人工知能は人間を超えるか」, 角川 Epub 選書, p46 では、「飛行機の場合は、鳥が飛ぶための「揚力」という概念を見つけ、揚力を得る方法(エンジンで推進力を得て、翼でそれを揚力に変える)を工学的に探索すればよかった。人工知能においても、知能の原理を見つけ、それをコンピュータで実現すればよい。」と説明している。
- (5)松尾豊,「人工知能は人間を超えるか」, 角川 Epub 選書, p61
- (6)「人工知能の未来 —ディープラーニングの先にあるもの—」(東京大学 松尾氏) ICT インテリジェント化影響評価検討会議 第1回 資料より

- (7)RAY KURZWEIL, The Singularity Is Near, 2005, Chapter3
- (8)Carl Benedikt Frey and Michael A. Osborne, THE FUTURE OF EMPLOYMENT: HOW SUSCEPTIBLE ARE JOBS TO COMPUTERISATION?, September 17 2013
- (9)論文中では, Paralegals, Contract and Patent lawyers が同じくくりで議論されており, Paralegals and Legal Assistants は 702 の職種中で 609 位 (上位ほどコンピュータへの置き換えが困難) という結果が示されている。一方で, legal assistants の入力したデータを活用する lawyers はコンピュータによる代替のリスクが低いと説明されており, 同ランキングでは 115 位という結果になっている。
- (10)新井紀子,「コンピュータが仕事を奪う」, 日本経済新聞社, 2010, p4
- (11)THE SECOND MACHINE AGE, p224
- (12)THE SECOND MACHINE AGE, p311
- (13)THE SECOND MACHINE AGE, p306
- (14)特許庁, 人工知能技術を活用した特許行政事務の高度化・効率化実証的研究事業 企画提案公募要領, 2016, URL: https://www.jpo.go.jp/koubo/koubo/pdf/jinkou_chinou/01.pdf (アクセス日: 2016 年 9 月 22 日)
- (15)特許情報プラットフォーム J-PlatPat による調査 (2016 年 9 月 22 日現在)
- (16)特許庁から FI・F ターム付与業務の委託を受けている一般財団法人工業所有権情報協力センターでは特許分類付与作業を効率化するために「つけつけ君 II」というシステムを用いている (参考: 古谷野浩志, 特許分類等の付与精度向上への取り組み, Japio 2007 YEAR BOOK, p118)。
- (17)FRONTEO 社ウェブサイト, URL: <http://www.kibit-platform.com/products/patent-explorer/> (アクセス日: 2016 年 9 月 23 日)
- (18)ヒューマンコンピューテーションについては, 鹿島久嗣, 小山聡, 馬場雪乃「ヒューマンコンピューテーションとクラウドソーシング (機械学習プロフェッショナルシリーズ)」, 講談社, 2016 の第 1 章から第 2 章に詳しく解説されている。
(原稿受領 2016. 10. 18)