

NEWSLETTER

知財速報

- 人工知能の波の下で、多くの作業が人工知能により置き換えられる。AIセマンティック技術は、特許検索において大規模に適用されているが、将来に検索エキスパートの強力なアシスタントになるだろうか、それともターミネーターになるだろうか。著者は、特許審査においてAIセマンティックを使用して検索を行った豊富な経験に基づき、AIセマンティック技術の生じた理由、基本原理、及び特許検索における適用方法に対する整理によって、答えを与えた。



AI セマンティック技術の特許検索における適用

人工知能の波の下で、多くの作業が人工知能により置き替えられる。AI セマンティック技術は、特許検索において大規模に適用されているが、将来に検索エキスパートの強力なアシスタントになるだろうか、それともターミネーターになるだろうか。著者は、特許審査においてAI セマンティックを使用して検索を行った豊富な経験に基づき、AI セマンティック技術の生じた理由、基本原理、及び特許検索における適用方法に対する整理によって、答えを与えた。

特許検索は 100 年以上の発展を遂げ、19 世紀の終わり頃、米国政府がコンサルティング業務において特許検索技術を適用しており、当時には紙の特許文献に対する手動検索であった。20 世紀 70 年代以降、コンピュータによる特許検索が徐々に普及され、これにより特許検索の効率も大幅に向上されたが、当時の一般人にとっては、特許情報を入手することは依然として非常に困難であった。1997 年、IBM がインターネットで特許情報サービスを提供して初めて、一般人の特許情報の入手が比較的簡単になってきた¹。特許検索の発展の歴史を見ると、情報技術の変革に伴い、特許情報の取得及び利用方法も絶えずに変化されてきた。

21 世紀に入って以来、情報技術の革新がますます加速しており、人工知能技術の発展が特に迅速となり、オンラインショッピングの時の推奨アルゴリズム、フェイススキャン、ボイスアシスタントなどの新技術は我々の日常生活に無意識のうちに溶け込んできた。特許検索分析の分野において、AI セマンティックと呼ばれる新技術も重要な発展期間に入ってきた。2018 年 5 月、世界知的所有権機関より「ICT 戦略及び人工知能についての知的財産局会議」が開催され、会議資

¹陳燕、黄迎燕、方建国。特許情報の収集及び分析[M]。清華大学出版社，2006：104-106。

料によると、複数の国や地域の知的財産局は、特許審査の検索プロセスにおいて AI セマンティック技術の適用を試し始めたそうだ²。商用特許データベースの分野では、多くのツールが既に AI セマンティック検索機能を備えていた。新技術の出現は特許検索業務にどのような変化をもたらせただろうか？著者は、特許審査において AI セマンティックを使用して検索を行ってきた豊富な経験に基づき、AI セマンティック技術の生じた理由、基本原理、及び特許検索における適用方法を簡単に紹介する。

従来の検索の直面したジレンマ

ジレンマは特許検索においてだけでなく、すべての従来のサーチ分野に存在している。情報の爆発的拡大がジレンマの主な原因であり、人間の知識が幾何学的な量で増えるにつれて、必要な情報を素早く正確にサーチされることは容易ではなかった。特許業界でも同じくデータの増大という問題に直面しており、中国本土だけで公開された特許文献の数、発明及び実用新案の特許文献の総数は、2008

² WIPO/IP/ITAI/GE/18, May 23 to May 25, 2018 (Geneva, Switzerland),

https://www.wipo.int/meetings/en/details.jsp?meeting_id=46586.

年末にはたった 250 万件未満だったが、2018 年末には 1700 万以上となり、10 年間で 7 倍増加して、これは特許検索業務にとって大きな課題をもたらせた。

従来のブール検索方式に基づいて、検索エキスパートらは検索の包括性と検索効率との間にバランスを見つけることに専心していた。いずれの技術的概念にもそれを表現する多くの方法があり、検索時にこれらの表現をできるだけ包括的に列挙してこそ、重要な文献を漏れないよう確保できる。しかし、各々の表現方式にはまた複数の意味を含んでおり、包括的な表現では不正確なノイズファイルを多く取り入れ、これらのノイズファイルは検索者の閲覧時間を無駄にした。今日、検索エキスパートらは何十倍も増加した特許文献に対し、検索の包括性と検索効率の間のバランスを見つけることがますます困難になっていると感じた。

AI セマンティック検索の台頭

従来の検索が直面したジレンマを解決するために、検索内容を自動的に拡張するセマンティック技術が出現され、具体的には、大きなシソーラスを構築し、シソーラスの中でユーザーの入力した内容をサーチした後、関連キーワードを拡張して検索のヒット率を向上させる。しかし、このようなセマンティック技術は、検索の包括性を向上させたが、このアルゴリズムの適用が毎回の検索でより多くの特許をヒットさせ、同じくノイズファイルを取り入れて、文献を閲覧及び選別する検索者の作業量を増加させた。

別の AI セマンティック技術は、セマンティック関連性のソートによるものである。このような技術は、特許テキストデータに基づいて機械学習が行い、語句間及び文書間の関

連性を自動的に演算し、関連性によって文献をソートする。このような方法の目的は、最も関連性の高い技術を最前に置いて、検索者による文献閲覧効率を直接に向上させる。この技術の特許検索ツールを適用すると、ユーザーが任意の長さの一段のテキストを入力したり、又は特許公開番号を直接入力したりすることが許可でき、システムが最も関連性の高い文献を自動的に推奨する。

AI セマンティック技術の特許検索における適用

シソーラスによるセマンティックサーチ技術は、コンピュータがユーザーの入力した内容を拡張しただけで、具体的な検索には依然としてブール検索を実行される。したがって、このようなセマンティック検索ツールは従来のブール検索ツールとの差がそれほど大きくなく、理解や使用上にてそれほど難しくない。

しかしながら、潜在的なセマンティック索引によるサーチ技術は、完全に従来のブール検索の範囲外であり、我々の検索に対する長期的な理解とは大きく異なる。このようなサーチ技術は、1 つの文献集合から部分集合を限定するプロセスを完全に放棄して、ユーザーの入力したテキスト内容の関連性に応じて、1 つの文献集合を直接ソートする。以下に、このような技術を適用して特許検索を行う 2 つの方法について簡単に紹介する。

第一の方法は、完全に独立して AI セマンティック検索を使用することである。このとき、検索者が検索しようとする一段のテキストをセマンティック検索システムに入力するだけで、コンピュータは、ユーザーの入力したテキストに応じてデータベースに格納されている何千万ないし何億もの特許についてソートすることができ、検索者は順番に閲覧す

るだけで必要な特許を取得することができる。このような検索方式は如何なる検索戦略も必要としないので、全く検索経験のない人であってもすぐに着手できる。検索経験が豊富である人にとっても、検索の初期段階でこのような方法を優先的に使用して探索的検索を実行することができ、素早く満足のいく結果が得られる一定の機会があり、かつ前にソートされた特許に対する統計分析によって、より多くのキーワードの表現方式又は関連する特許分類番号を見つけることができる。

完全に独立して AI セマンティック検索を使用することは、検索効率を向上させることはできるが、まだ検索エキスパートやブール検索を置き換えることはできない。その主な理由の一つは、AI セマンティック技術が技術手段方面の理解を行うことができなく、つまり、AI セマンティック技術は、まだ検索エキスパートのように、正確な検索を実行するための技術手段におけるコア技術的特徴を抽出することができない。例えば、以下の一段のテキストについて、検索エキスパートは、そのうちに記載されている技術手段及び従来技術に対する理解に基づいて、「切り欠き」や「折曲部」の2つの特徴をコア特徴として検索を実行するが、現在の AI セマンティックアルゴリズムは、まだ検索エキスパートのように、テキストから技術手段までの抽象的な思考プロセスをシミュレートすることができない。

伸縮ロッド、及び、ステージとステージの上に設けられる伸縮式クランプ機構とを含む、撮像装置を保持するためのクランプ装置、を備える一体型セルフタイマー装置において、前記クランプ装置は、前記伸縮ロッドの先端に一体的に回転接続され、前記ステージの上には、切り欠きが設けられ、前記クランプ機構には、前記切り欠きの位置に対応した折曲部が設けられ、前記伸縮ロッドは、折り畳まれた後に前記切り欠き及び折曲部に収容されることがで

きることを特徴とする一体型セルフタイマー装置。

AI セマンティックを使用する最善の方式は、エキスパートブール検索と組み合わせて用いることである。具体的には、ブール検索式により検索結果が得られた後、さらに1つのブール検索式とは無関係のソート因子により前述の検索結果をソートする。ブール検索式は、技術手段のコア特徴を正確的に表現するためのものであり、AI セマンティックは、他の非コア特徴を関連性で表現し、ソートする方法によって検索エキスパートに提示させるものである。このような検索の利点は、一方では検索エキスパートのコア特徴を定義する役割を十分に発揮して、検索の精度が確保できると同時に、AI セマンティックアルゴリズムのサーチの包括性を発揮して、非コア技術的特徴の表現の不十分による検索漏れを避けることができる。

上記の例では、ブール検索を単独で用いる場合、少なくとも検索時に、セルフタイマー装置、クランプ装置、切り欠き、折曲部などの特徴が限定される必要があるが、セマンティックとブール検索を組み合わせる場合には、ブール検索のみを利用して切り欠き、折曲部などのコア特徴をサーチし、またこのテキスト全体をソート因子として検索結果をソートすることにより、セルフタイマー装置、クランプ装置に関連する特許を最前にソートすることができる。このような検索方式は、検索の正確さを向上させるだけでなく、重要な特許の検索漏れも回避して、検索の効率も大幅に向上させることができる。

人工知能がいくつかの仕事を取り替えてくる話しを聞く度に、少なくない人が少々心配してしまうと思う。しかしながら、筆者は、特許検索自体が、その目標の高さが不確実で、多くの主観的思考、判断及び総合的分析を必要とする作業であると考えられており、検索

エキスパートらは、AI セマンティックを競争相手として見なす必要がなく、将来の作業において自分自身に力を与えるための重要なツールと考えて良い。セマンティック技術とエキスパートブール検索との組み合わせが将来

の検索の主流になると考えられており、大量のデータモデルによるセマンティックソート及びブール検索による正確なヒットは、人間と人工知能の融合による完璧なアンサンブルになるだろう。

この文章は法律意見書と同等ではありません。具体的な法律意見書については、当社の専門コンサルタントや弁護士にご相談ください。当社の電子メールは LTBJ@lungtin.com、当該電子メールは当社のウェブサイト www.lungtin.com でも見つけます。

詳細な情報やさらなる助言については、この文章の筆者にお問い合わせください。

周俊：シニアプロジェクトマネージャー、特許アナリスト、元特許審査官：LTBJ@lungtin.com



周俊

(シニアプロジェクトマネージャー、
特許アナリスト、元特許審査官)

周俊先生は、大連理工大学車両工学の修士、中国国家知識産権局の元特許審査官で、国家知識産権局知的財産発展研究センター(CNIPA-IPDRC)で臨時任職されて、特許分析プロジェクトの管理及び研究作業に参加しており、現任中専隆天知識産権運営(深セン)股份有限公司のシニアプロジェクトマネージャーである。周俊先生は、国家知識産権局が主催した複数の特許早期警戒及び特許ナビプロジェクトに参加して、国家知識産権局及び各級政府の特許と産業戦略の策定に支援しており、複数の企業向けの特許競争性情報分析プロジェクトを主導して、企業クライアント様の研究開発の方向性判断、競争他社の技術動向の理解、技術競争のパターンの明確、及び革新的なリソース統合などに助けを与えた。彼は新エネルギー車やインテリジェントネットワーク車分野に取り込み、特許支援研究開発及び特許ポートフォリオ戦略などの方面で豊富な知識と経験を持つ。