

特許庁におけるAI活用の現状とマルチモーダル技術の産業財産権制度への影響について

The Current State of AI Application in JPO and the Impact of Multimodal Technology on the Industrial Property Rights System

特許庁 総務部総務課 情報技術革新室長

櫛引 明佳

2002 年特許庁入庁。審査第三部で化学分野を中心に審査業務に従事するとともに、外部機関への出向、ユーザ側（審判課、商標課・意匠課）とシステム開発室で複数のシステム刷新プロジェクトに参画。2023 年 7 月より現職。

特許庁 審査第二部生産機械（工作機械）技術担当主任
（前：特許庁 審査第一部調整課審査企画室 企画調査官）

五十嵐 康弘

2005 年特許庁入庁。審査第二部で審査業務に従事するとともに、ユーザ部門、開発部門の双方でシステムに携わる。2024 年 1 月より特許審査における生成 AI の活用検討に携わる。

1 はじめに

本論では、まず、特許庁における AI 活用の現状と今後の方向性について紹介したのち、本特集のテーマである「マルチモーダル技術が知財情報にもたらす未来」について考察していきたい。本論は、調整課五十嵐企画調査官との共著とさせて頂き、特許庁における AI 活用の全般的な取組や商標領域については小職が、特許領域については五十嵐氏に執筆頂いた。

特許庁における AI 活用の現状は、テキスト情報や画像情報のそれぞれを中心とした「シングルモーダル」に関するものが中心であって、近年生成 AI 等を中心にマルチモーダル技術が著しい発展を遂げているといえども、直ちに産業財産権の保護制度そのものに影響を及ぼすものとは考えにくい。

とはいえ、このようなマルチモーダル技術の発展が将来的に審査官の検索・判断行動やユーザである各法域の出願人の保護対象に関するニーズに変化をもたらす可能性があることも容易に想像できるところであり、本論では、筆者の個人的見解や想像に基づく自由な見解を述べさせて頂くことをあらかじめお断りしておく。

2 特許庁における AI 活用の現状

2.1 近年のアクション・プランの変遷

特許庁では、2017 年に、特許行政事務の高度化・効率化に向けて「人工知能（AI）技術の活用に向けたアクション・プラン」（以下、「アクション・プラン」）を策定し、その推進のために人工知能関連技術活用可能性検証プロジェクトチーム（AI 検証 PT）及び内製アジャイル開発の体制を併せて立ち上げ、かかる枠組みのもと AI 活用の取組を継続してきた。

現状のアクション・プランについては、2022 年（令和 4 年）5 月に公表した「アクション・プラン（令和 4～8 年度版）」をベースに、問い合わせ対応等の特許行政事務への生成 AI の適用等、令和 6 年度に実施する技術実証に関する取組を掲載した「アクション・プラン（令和 6 年度改定版）」を経て、これまでの検討結果を踏まえるとともに、生成 AI の活用をさらに進めるべく見直しを行い、2025 年 6 月に公表した「アクション・プラン（令和 7 年度改定版）」（図 1）が最新である。

令和 7 年度改定版の改定の観点としては、1）今年度までの実証事業の結果を踏まえ、「3. 先行技術調査②」（検索手法の高度化）について導入フェーズへ移行することとした点、2）令和 6 年度改訂版における「4. 特

「特許審査管理業務」について、過去の実証事業を踏まえた検討の結果、当該実証事業による検証は一時凍結する一方で、「9. 生成 AI の特許審査業務への適用」を新設して、生成 AI の活用を検討していくこととした点、3) 2025 年（令和 7 年度）に「3. 先行技術調査②」、「6. 指定商品・役務調査」、「8. 生成 AI の特許行政事務への適用」、「9. 生成 AI の特許審査業務への適用」について技術実証を実施することとした点、の 3 つの観点を反映させたものである。

2.2 アクション・プランの今後

現行アクション・プランは令和 4 年度から令和 8 年度までの 5 ヶ年の計画を定めたものであるが、2026 年（令和 8 年）以降の計画については現在未定である。

また、現行アクション・プランの立案時には現在の生成 AI の発展は想定していなかったものであり、生成 AI が様々な業務領域で活用されていることからすれば、実体審査業務はもちろんのこと、その他の特許行政事務業務全般において生成 AI の活用により業務改善ができる可能性を大いに秘めており、当面の検討の中心は生成 AI に移っていくことになると思料する。詳細は五十嵐氏に譲るとして、マルチモーダル技術の発達と生成 AI の登場により、特許領域で主に扱われてきたテキストや図面に加え、音声や動画等の情報がこれまでとは比較にならないほど容易に製造できることとなり、それは同時にそのような検索が可能になることと同義であるし、特許だけに限らず他の法域においてもマルチモーダル技術の審査実務への影響が及ぶことになる。現在、AI 技術の特許庁業務への活用に取り組む筆者にとっても、マルチモーダル技術は生成 AI と併せて非常に興味深い技術である。（櫛引）

3 特許審査におけるマルチモーダル技術が及ぼす影響

3.1 マルチモーダル情報

近年、AI は、画像認識、翻訳、自然言語理解、生成 AI など多方面で飛躍的な進化が進んでいる。

マルチモーダル情報を扱う大規模言語モデル（以下、MLLM）とは、複数の異なる種類のモード（情報種別）を統合して扱う AI 技術である。例えば、テキストと画像、テキストと動画などを同時に処理し、AI が対象の情報

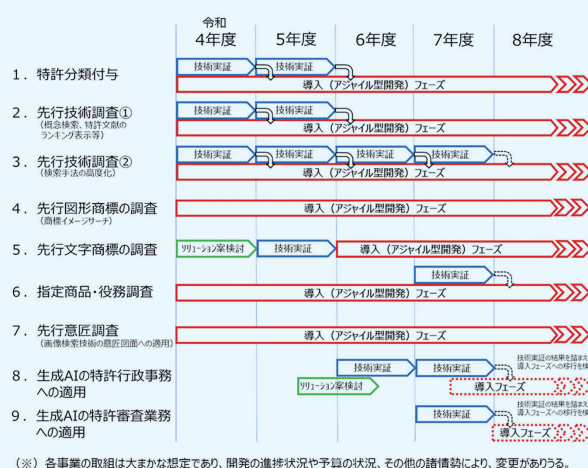


図1 アクション・プラン（令和7年度改訂版）

をより高度に解釈し、出力を行うことができるとされる。他にも様々なセンシング情報を合わせると、人間の五感では感知できない情報をも包括的に扱える可能性も秘めている。

一方、人間は、特段意識せずとも、周囲の環境から視覚、聴覚などを通して情報を得ており、マルチモーダル情報を処理している。

3.2 人と人とのコミュニケーション

MLLM は比較的新しい技術であることから、まずは、従来からある人と人とのコミュニケーションで扱う情報種別について考えてみる。

言語は単語と文法の組み合わせにより多様な情報を表現できる。人と人のコミュニケーションにおいて、古くは会話による直接の伝達に始まり、有史以来は文字をもって時と場所を超えた伝達が行なわれ、近代では発声された情報も記録、伝送されるようになった。このように、言語は、様々な情報種別の中でも特別な位置を占め、コミュニケーションの軸をなしている。

人と人との会話では言語の発声に加えて、身振り、無意識の仕種や目線、表情、声色など多くの情報を添えてコミュニケーションを行う。目は口ほどにものをいうといわれるように、言語以外の情報の重要さは昔から知られていた。

とはいえ、近年、テレワークの浸透によっても言語以外の情報の価値が再認識され、身をもって利点を感じている人は多いだろう。

他方、メールやチャットなどテキストベースのコミュニケーションでは、絵文字やスタンプの利用、写真やス



クリーンショットの添付、返信の間隔など、これも多くの情報を添えて相手に伝えたい内容を調整する。このように、人と人のコミュニケーションについてみると、言語を軸としつつ、言語以外の情報種別も加えて、すなわちマルチモーダル情報を用いて豊富な情報をやりとりしている。

3.3 マルチモーダル情報の2つの作用

マルチモーダル情報についてみると、言語以外の情報種別は、大きく2つの役割を果たしているように思われる。

一つは、言語のみでは表現できない情報を付加したり、相手が受け取る言語情報の解釈の余地を限定したり、解釈を誘導したりするように作用し、実質的な意味を付加、特定する情報である（以下、「意味特定のマルチモーダル情報」という。）。

もう一つは、相手の認知負荷を軽減し、理解を容易にするように作用する情報である（以下、「理解補助的なマルチモーダル情報」という。）。

この2つは、必ずしも厳密に分けられるものではなく、両方の作用を持つ情報もあるが、より正確に情報を伝える、より円滑に情報を伝える、という2軸があるといえる。

3.4 特許出願に含まれるマルチモーダル情報

さて、ここで特許出願に含まれる情報について考えてみよう。

特許出願は、願書に明細書、特許請求の範囲、必要な図面、要約を添付して出願され（特許法第36条）、一部には図や表による情報が含まれるものの、多くはテキストとして表現される。

このテキストは、発明の内容を公開する役割（第64条）であったり、特許発明の技術的範囲を確定させる役割（第70条）であったり、特殊な目的で作成される。

そのため、ほかの目的で作成されるテキストと比べると、その技術分野の専門家（第36条第4項）を水準として読まれることを前提として、ゆらぎなく意味を解釈できるように作られる。また、手続の補正を行う場合の補正できる範囲の制約となる（第17条の2第3項）。その結果、ほかのテキストに比べて、十分な記載内容、正確な記載内容となるように、例えば同様の記述の繰り返しが多くなるなどの弊害があったとしても、冗長に記述され、数百ページを超える特許出願も少なくない。

図面（図面として添付される場合のほか、特許請求の範囲や明細書に挿入される図も含む）はピクセルによって情報が構成され、必要に応じて表、グラフ、化学構造式、回路図、フローチャート、機器や部品の形状など、多様な情報を表現するように描かれる。

しかし、上述の通り、特許出願はテキストで正確に記述されるため、発明の主要な部分を特定するための唯一の手掛かりとして、図面の情報が用いられることは少ない（化学構造式などが例外にあたるだろうか。）。テキスト以外の情報は、それを読む人の理解を促すように用いられる、すなわち、理解補助的なマルチモーダル情報であることが多い。

3.5 検索のためのシングルモード化

まず、特許審査において意味特定のマルチモーダル情報をどう活用しうるかを考えてみる。

特許審査で審査対象となる要件は多数ある（第49条）が、先行技術調査を要するいわゆる新規性、進歩性の判断は特許審査のタスクの多くを占める。新規性、進歩性以外の要件については、明細書等の情報に基づいて判断が可能な場合が多いのに対して、新規性、進歩性の判断は先行技術（第29条第1項各号）に該当する膨大な情報を考慮して判断する必要があるためだ。

先行技術調査は、特許文献や学術論文等を中心に行われるが、特に特許文献以外の情報は、論文、書籍、雑誌などの形式で存在し、想定している読み手も様々である。学術論文では、情報の正確さや十分さは重視されるが、上述のような特許文献独特の冗長性の弊害を容認したように記載されているとは限らず、図やグラフが主要な部分を表すことは多々ある。

さらに、近年、動画での情報発信が急激に増えている。それに加え、人々に提供される視覚情報が飽和しつつあるせいか、ポッドキャスト等の音声メディアもラジオ最盛期以来の復興を見せつつある。

動画や音声といった情報は、視覚、聴覚に作用するが、時系列データであるという特徴がある。すなわち、先行技術調査として動画や音声といった情報を調査しようとすると、必然的にそれらの視聴が必要になり、膨大な時間を要することとなる。

また、動画や音声情報を提供するプラットフォームでは、タイトル、概要文や、視聴者からのコメントなど多くのテ

キスト情報も併せてマルチモーダル情報を構成する。

現状、特許審査における動画の引用頻度は高くない。しかしながら、学会講演の録画や、製品の宣伝動画、科学技術の解説動画など、着実に先行技術となりうる動画、音声情報は増加している。今後、それらが示す情報を検索する必要は、増加の一途を辿るであろう。

しかしながら、先行技術調査は、特許分類やテキスト検索、単語頻度に基づく類似度などの手法を用いており、動画、音声といった情報を直接検索することができない。

そこで、意味特定のマルチモーダル情報から、図やグラフ、時系列データと共にテキスト情報なども考慮した上で、テキストに変換する用途で MLLM が役立つと考えられる。時系列の要素が排除されるように変換、蓄積することで、調査に要する時間は激減し、従来の学術文献等と同様に調査を行うことができるようになる。

もちろん、最終的には MLLM で変換する前のマルチモーダルな一次情報に基づいて先行技術を認定し、審査を進めるが、多数の情報から取捨する作業に比べれば現実的なタスクとなる。

3.6 理解補助のためのマルチモード化

続いて、理解補助的なマルチモーダル情報についても検討してみる。

特許審査において、最初に審査官が行うべきタスクは、出願された発明の理解である。明細書、特許請求の範囲、図面等を読み込みながら理解を進めていくが、上述のとおり、明細書等は正確な情報が記載されているものの、通読でスムーズに内容を理解できるように、といった点が考慮されているとは限らない。とりわけ、外国特許庁の出願を優先権基礎とする出願は、技術事項の記載順序などに関して、国ごとの特徴もある。その結果、日本語に翻訳された文であっても、何度も何度も読み返しながら、徐々に理解を進める必要がある出願も多い。

そこで、マルチモーダルな情報を出力できる MLLM を用いて、明細書等の情報を入力として、理解しやすさに特化して順序だてられた内容に組み換え、必要であれば理解しやすい図も添えることで、審査実務において大いに役立つと考えられる。また、検索によって見つかった先行技術についても、審査官は精読し内容を理解していく必要がある。その際にも、理解しやすさに特化した加工ができれば、これも有用であろう。

人同士が、身振り手振りを添えながらコミュニケーションを取るように、人と MLLM とが、理解しやすいように情報種別を変えながらやり取りをできれば、適切な技術情報を探し、それ容易に理解できる未来が描ける。

それは、単に特許審査の場のみならず、情報をより多くの人に、親しみやすく広めることとなり、知が尊重される社会の実現に繋がるのではないだろうか。(五十嵐)

4 商標審査におけるマルチモーダル技術が及ぼす影響

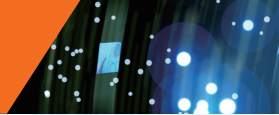
特許審査官である筆者が商標制度について見解を述べるのは大変恐縮であるが、AI の業務領域への適用推進に従事する筆者の見解としても、マルチモーダル技術は商標制度に大きな影響を与え得る可能性を秘めた技術であると考ええる。

4.1 保護対象となる「標章」の多様性とマルチモーダル技術の関係

マルチモーダル技術と商標の関係について、新しいタイプの商標制度として、動き商標、ホログラム商標、色彩のみからなる商標、音商標、位置商標が保護の対象として認められており、特許が基本的にテキストと図面で出願内容が構成されていることと比較すると、動き商標やホログラム商標のような図形そのものに加えて動的な変化を含む標章が保護対象として既に認められているし、音商標に至っては聴覚で知覚される情報である。また、新しいタイプの商標制度法改正時には保護対象とはならなかったが、「におい」や「味」についても候補として検討されていたように、五感で認知可能な情報またはそれらを組み合わせで構成された表現物を保護対象とするニーズがそもそも存在している。

そして、マルチモーダル技術と AI の関係については五十嵐氏が述べたとおりであるところ、生成 AI による創作と商標制度との関係については、本年6月の産業構造審議会知的財産分科会第12回商標制度小委員会において、人が AI を利用して生成した商標や AI が自律的に生成した商標であっても、商標法第2条の定義を満たすのであれば、「商標」に該当し、保護の対象となることが示されたことから、今後は AI を活用した「商標」の創作が進んでいく可能性は高いと考えられる。

今後のマルチモーダル AI の発達により、視覚情報に



加えて、聴覚、触覚、味覚、嗅覚などを複合して構成される複雑なモダリティから構成する標章の創作が容易になっていくことは明らかであり、そのような標章が実社会で識別力を有してブランドとして定着していくことになれば、新たなタイプの商標としての保護の必要性がさらに高まってくることは十分に考えられるのではないか。

4.2 審査実務に及ぼす影響

かりに、このような複雑な商標出願のニーズが高まった場合に、実務上どのような対応が必要になるだろうか。

まずは、より複雑な商標出願に対応できる出願制度の整備が求められる。現制度下においては、新しいタイプの商標であっても、出願時には原則「商標の詳細な説明」において保護を受けようとする商標を文言で特定する必要があるが、例えば、色彩のみからなる商標においては、色彩の組み合わせをRGBの配合率や色見本帳の番号で特定することができる。そうすると、マルチモーダル技術の活用によって、複雑な商標の特徴を規格として確立された数値等に変換することができるのであれば、保護の対象とし得る可能性が出てくるかもしれない。また、物件での提出を認めるとしても、複数のモダリティを正しく表す出願方式（電子データの形式）の整備が求められることになる。理論上可能とも思われるが、実現に向けたハードルは高いであろう。

そして、複雑な特徴を有する商標を検索するための検索システムの整備が必要になるが、すでに画像検索システム等の実装が進んでいることを考えると、これについては比較的現実味があるように思える。

実務上は、このような商標に対する実体審査を行うとすると、商標を構成する個別のモダリティの類似性判断に加えて、それらが複合的に組み合わせられた際に何をもちて類似・非類似とするのか新たな判断基準の整備が当然必要になってくるうえ、個々のモダリティの類似性を判断する際においても、かりに特徴が数値化されて出願されたものであるとすると、その数値の技術的な意味合いを理解する必要がある可能性が高いと考えられ、審査官、弁理士の双方に求められる知識もより高度なものになっていくかもしれない。

無論、現実的には、上述のように、対象を適切に保護できる法制度整備、より複雑な識別力や類似性の判断手法の確立、国際制度調和、専門的知見を有する弁理士や審査官

等の人材育成等の多くの課題が存在すると考えられるから、マルチモーダル技術が進歩したといえども、短期間の間に制度が大きく変化することはないと思われる。

以上、制度面の細かい点は深く立ち入らずに、技術面から述べさせて頂いた。意匠制度については触れなかったものの、基本的には商標の場合と同様の変化の方向性になるであろう。（櫛引）

マルチモーダル技術が知財情報にもたらす未来