

AI関連発明の特許出願状況と日本特許庁における取り扱い

弁理士 黒木 義樹



1 はじめに

近年、深層学習(ディープラーニング)を中心とする機械学習を利用した人工知能(AI; Artificial Intelligence)技術が急速に普及しており、第3次AIブームと言われています。IoTデバイスによるデータ収集の容易化、ディープラーニング技術の進化、およびハードウェア性能の向上という三要素が重なった

ことがブームのベースにあります。これに伴い、AI関連の特許出願も急増しており、IT企業だけでなく、自動車、機械、金融、化学、材料、医療など技術分野をまたがって出願が増加しています。

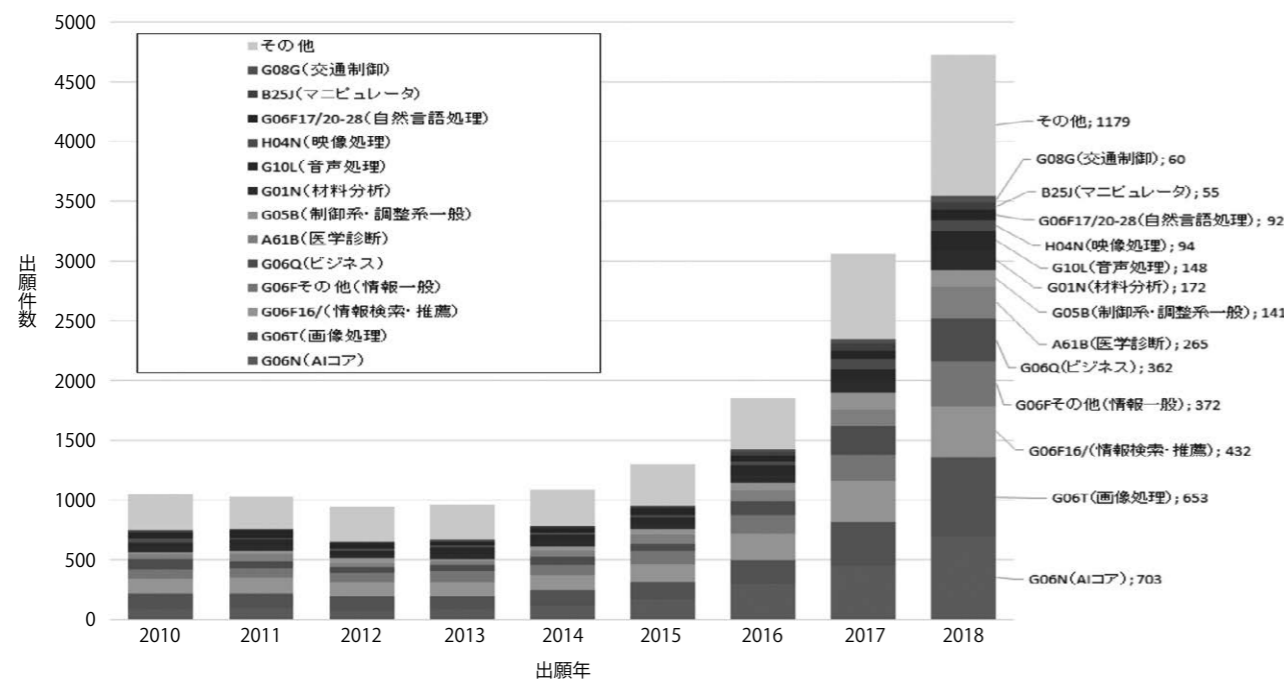
この状況を受け、日本特許庁は2019年7月に「AI関連発明の出願状況調査報告書」を公表し、このたび2020年7月に同報告書を更新しました。

本稿では、更新された同報告書に基

づいてAI関連発明の特許出願の状況を報告するとともに、日本特許庁における取り扱いについてお話しします。

2 AI関連の特許出願の状況

下記グラフは、AI関連発明の出願件数について、主分類ごとにその推移を示しています。ここでAIについては様々な定義が可能ですので、①AIコア発明及び②AI適用発明を「AI関連発明」と定義し、調査対象とされています。



AI関連発明の主分類構成の推移 (2018年の件数を表記)¹

と定義し、調査対象とされています。

AIコア発明とは、ニューラルネットワーク、深層学習などの各種機械学習技術のほか、知識ベースモデルなどAIの基礎となる数学的又は統計的な情報処理技術に特徴を有する発明です。

AI適用発明とは、画像処理、音声処理、機器制御・ロボティクス、診断・検知・予測・最適化システムなどの各種技術にAIコア発明を適用したことを特徴とする発明です。

下記グラフに示されるように、AI関連発明の出願件数は2015年(出願年)ごろから増加傾向にあり、2018年(出願年)には約4700件となっています。内訳を見ると、AIコアの分類であるG06Nが付された出願以外では、主分類としてG06T(画像処理技術)やG06F16/(情報検索・推薦)が付与された出願が多く見られます。

G06Fその他(情報一般)の出願も多く、これにはG06F3/(インターフェイス)やG06F21/(セキュリティ)といったAIの適用先が含まれます。

他にも、G06Q(ビジネス)、A61B(医学診断)、G05B(制御系・調整系一般)、G01N(材料分析)、G10L(音声処理)、H04N(映像処理)などもAIの主要な適用先になっています。

このように、AI関連発明の特許出願は増加を続けており、またAIの適用先も幅広い技術分野に広がっていることが分かります。

次項では、AI関連発明が日本特許庁においてどのように取り扱われるかお話しします。

3 AI関連発明の取り扱い

AI関連発明の審査は、現行の審査基準等に基づいて特段問題なく行えているというのが、日本特許庁の現状のスタ

タンスです。

しかし、今後もAIが様々な技術分野に適用されて発展することが予想されているため、AI関連発明と関連が少なかつた出願人等に審査の運用を分かり易く示し、また特許庁としても技術分野によらず統一的に特許性の判断を行う必要があるため、事例集²に特に問題となる発明該当性、記載要件、及び進歩性の事例を集積しています。

(1) 発明該当性

機械学習においては、まずは「生データ」が必要です。生データとしては、テキストデータや画像/音声データ、計測データなどが挙げられます。生データは、学習プログラムによる学習に適するように相関の高い特徴量を抽出したり、正解をラベルしたりして調整され、より価値の高い「学習用データ」が生成されます。

収集されたこれらデータは一定の価値を有するため、法的保護が必要な場合がありますが、生データも学習用データも情報の提示を主目的とするものですので、技術的思想の創作に該当せず発明該当性が否定され、特許法での保護は難しいです(不競法や著作権法等での保護の可能性があります)。この点、同事例集の[事例3-2]では、「反射式近赤外分光分析を行う携帯型のリンゴ用糖度センサにより計測された、果樹に実った収穫前のリンゴの糖度データ」のクレームが、発明に該当しない事例として紹介されています。

ただし、学習用データについては、いかにして学習に使用されるかという観点から構成を限定するものであれば、「データ構造」のクレームとして特許の可能性がります。この点、同事例集の

[事例2-13]には、必ずしも学習用データに直接かかわるものではないものの、「音声対話システムの対話シナリオのデータ構造」が発明に該当する事例として紹介されています。ポイントは、「データ構造」がプログラムに準ずるもの、すなわち、データの有する構造がコンピュータの処理を規定するものという点でプログラムに類似する性質を有するものであれば、ソフトウェア(プログラムに準ずるデータ構造)とハードウェア資源とが協働した具体的手段又は具体的手順によって、使用目的に応じた特有の情報の演算又は加工が実現されているとして、発明該当性が認められる可能性があります。また、学習用データの作り方に特徴があれば、「学習用データの生成方法」としても権利化の可能性がります。

「機械学習アルゴリズム」については、新規なアルゴリズムを発明したとしても、関数やニューラルネットワーク等の数理的なモデルを単に示すものであれば、自然法則を利用するものではなく発明該当性が否定される可能性が高いです。しかし、アルゴリズムはソフトウェアにより具現化されるものですので、コンピュータソフトウェア関連発明と捉えることができるようにクレームすれば、装置や方法、プログラムの発明として発明該当性が認められる可能性があります。

「学習済みモデル」は、学習段階で得られた関数に入出力を絡めて推測段階の全体を学習済みモデルとして捉えれば、発明該当性が認められる可能性が高まります。同事例集の[事例2-14]には、宿泊施設の評判を分析するための学習済みモデルの発明が紹介されています。なお、この事例ではクレームでニューラルネットワークの構成まで限

定していますが、この部分をブラックボックスとし、入出力データによる限定により特許されている出願も散見されます。入出力データの相関が未知であれば(既知なら進歩性の点で厳しい)、権利行使時の立証の容易さを考えると、後者でまずはチャレンジしたいところです。

最後に、学習済みモデルを搭載した機器やシステムとしての「AIシステム」は、通常のソフトウェア関連発明と同様に、機器等に対する制御又は制御に伴う処理を具体的にを行うもの、又は対象の物理的性質又は技術的性質に基づく情報処理を具体的にを行うものであれば、発明該当性が認められる可能性が高いです。一方で、ビジネスモデルなどの非技術なシステムの場合は、ハードウェア資源を用いて具体的に実現されていれば、発明該当性が認められる可能性が高いです。

(2) 記載要件

特許出願の審査では記載要件として実施可能要件・サポート要件などが審査されますが、AIを様々な技術分野に適用したAI適用発明では、ある程度の推定精度を有する学習済みモデルが生成可能なことが十分に説明されていなければ、記載要件違反となってしまう。

ディープラーニングのようなデータに基づき統計・確率的な手法を用いたデータ駆動型AIでは、判断処理を行う部分がブラックボックスであるため、学習済みモデル生成に用いる学習用データ(教師データ)に特徴を持たせて出願することが多くなります。しかし、実現性のない発明や、実態からかけ離れてクレームが広すぎる発明に特許を付与するのは適切ではありません。

そこで特許庁は、学習用データ(教師データ)の「相関関係等」に注目して記載要件を判断しています。同事例集に

は、次の通り示されています。

AIを様々な技術分野に応用した発明は、AIの機械学習に複数種類のデータを含む教師データを用いることが一般的であるが、この場合、記載要件を満たすか否かの判断において、発明の詳細な説明の記載に基づいて、当該複数種類のデータ間に相関関係等の一定の関係(以下、「相関関係等」という。)が存在することが認められること、又は、技術常識に鑑みて当該複数種類のデータ間に何らかの相関関係等の存在を推認できることが必要である。しかし、発明の詳細な説明に、複数種類のデータ間の具体的な相関関係等が開示されている必要はない。

そして、同事例集では、次の3つの観点から記載要件を判断するとしています。

- a) 出願時の技術常識を鑑みて教師データに含まれる複数種類のデータ間に相関関係等が存在することが推認できる
- b) 教師データに含まれる複数種類のデータ間の相関関係等が明細書等に記載された説明や統計情報に裏付けられている
- c) 教師データに含まれる複数種類のデータ間の相関関係等が実際に作成した人工知能モデルの性能評価により裏付けられている

同事例集の[事例46]には、人物の顔画像を入力し、その人物が栽培した野菜の糖度を推定する糖度推定システムの発明が紹介されています。しかし、出願時の技術常識に鑑みても、人物の顔画像とその人物が栽培した野菜の糖度との間に相関関係等の一定の関係が存在することは推認できず、また実際に生成された判定モデルの性能評価結果も示されていないとして、実施可能要件違反とされています。

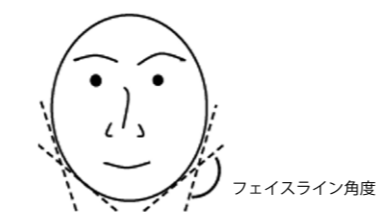
一方[事例47]では、商品のウェブ上での広告活動データ及び言及データを入力し、その商品の今後の売上数を予測することをポイントとする事業計画支援装置の発明が紹介されています。

この事例では、出願時の技術常識に鑑みてウェブ上での広告活動データ及び言及データと、売上数との間に相関関係等の一定の関係が存在することは推認できるとして、実施可能要件を満たすとされています。

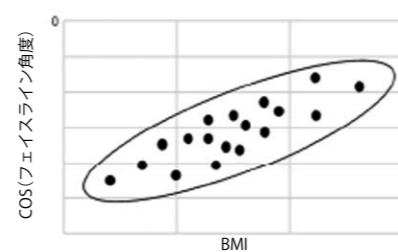
また[事例49]では、人物の顔の形状を表現する特徴量及び身長を入力し、体重を推定する体重推定システムの下記発明が紹介されています。

【請求項1】
人物の顔の形状を表現する特徴量と身長及び体重の実測値を教師データとして用い、人物の顔の形状を表現する特徴量及び身長から、当該人物の体重を推定する推定モデルを機械学習により生成するモデル生成手段と、人物の顔画像と身長を入力を受け付ける受付手段と、前記受付手段が受け付けた前記人物の顔画像を解析して前記人物の顔の形状を表現する特徴量を取得する特徴量取得手段と、前記モデル生成手段により生成された推定モデルを用いて、前記特徴量取得手段が取得した前記人物の顔の形状を表現する特徴量と前記受付手段が受け付けた身長から体重の推定値を出力する処理手段と、を備える体重推定システム。
【請求項2】
前記顔の形状を表現する特徴量は、フェイスライン角度であることを特徴とする、請求項1に記載の体重推定システム。

〈図1〉



〈図2〉



この事例では、頬のラインとあごのラインが形作る角度をフェイスライ

ン角度と定義すると〈図1〉、「フェイスライン角度の余弦」と「その人物のBMI(体重/(身長²))」との間に統計的に優位な相関関係があることを見出した〈図2〉ことから、人物の顔画像を解析して、その顔の形状を表現する特徴量と身長から体重を推定しています。

この事例では、出願時の技術常識からはフェイスライン角度等の顔の形状の特徴量と、その人物の身長及び体重やそれらに基づくBMIとの間に相関関係等は推認できないものの、フェイスライン角度の余弦とBMIとの間に統計的に有意な相関関係が存在することが明細書に示されているとして、請求項2については記載要件を満たすとされています。一方、フェイスライン角度以外の顔の形状の特徴量との具体的な相関関係については示されておらず、また推定モデルの性能評価結果も示されていないとして、請求項1については実施可能要件違反及びサポート要件違反とされています。

また[事例50]では、被験物質のヒトにおけるアレルギー発症率を予測する方法の下記発明が紹介されています。

【請求項1】
ヒトにおけるアレルギー発症率が既知である複数の物質を個別に培養液に添加したヒトX細胞の形状変化を示すデータ群と、前記既存物質ごとのヒトにおける既知のアレルギー発症率スコアリングデータとを学習データとして人工知能モデルに入力し、人工知能モデルに学習させる工程と、被験物質を培養液に添加したヒトX細胞において測定されたヒトX細胞の形状変化を示すデータ群を取得する工程と、学習済みの前記人工知能モデルに対して、被験物質を培養液に添加したヒトX細胞において測定されたヒトX細胞の形状変化を示す前記データ群を入力する工程と、学習済みの前記人工知能モデルにヒトにおけるアレルギー発症率スコアリングデータを算出させる工程とを含む、ヒトにおける被験物質のアレルギー発症率の予測方法。
【請求項2】
ヒトX細胞の形状変化を示すデータ群が、ヒトX細胞の楕円形度、凹凸度、及び扁平率の形状変化の組合せであり、アレルギーが接触性皮膚炎である、請求項1に記載の予測方法。

この事例では、出願時の技術常識からはアレルギー発症率と細胞の形状の変化との間に相関関係等の一定の関係が存在することは推認できないものの、ヒトX細胞の形状変化として、楕円形度、凹凸度、及び扁平率の組合せについては、人工知能モデルの性能評価により一定精度でアレルギー発症率を予測できたことが明細書で裏付けられているとして、請求項2については記載要件を満たすとされています。一方、上記3種以外の形状変化の組合せについては、具体的な相関関係については明細書に示されておらず、また推定モデルの性能評価結果も示されておらず、さらに接触性皮膚炎以外のアレルギーについての予測可能といえる合理的な理由もないとして、請求項1については実施可能要件違反及びサポート要件違反とされています。

このように、AI関連発明の出願では、学習用データに相関関係等の一定の関係が推認できない場合は、明細書中で十分な説明や統計情報による裏付け、さらにはAIモデルの性能評価の提示などが必要となるため、注意が必要です。

(3) 進歩性

進歩性については、今あるものに単純にAIを適用したというだけでは認められ難いです。これに対し、入出力データとして各技術分野で未知のパラメータを入れて特徴を出せば、進歩性が認められる可能性が高まります。

[事例34]には、ダム水力発電量を高精度に直接推定できる水力発電量推定システムの発明が紹介されています。この事例では、入力データとして「上流域の降水量」、「上流河川の流量」及び「ダムへの流入量」をパラメータとして使うことは公知であるものの、「上流域の気温」という未知のパラメータをさらに使用することで、より高精度

な発電量の推測ができるという有利な効果を奏し、進歩性が認められています。

AIの分野において先行技術文献の蓄積が十分に進んでいない現状では、未知のパラメータを規定して特許になることも多く、AI関連発明はいち早く出願する意義が高いと思います。

なお、進歩性については、次回以降の記事でも改めて取り上げたいと思います。

4 その他

AI関連発明においても、審査で特許にすることだけを考えるのではなく、権利行使の場面を意識して請求項を作成する必要があります。クラウドを介してサービスが提供される場面も多いと思われるところ、侵害行為を常に意識し、一当事者の行為に注目して、侵害発見が容易な請求項を作成する必要があります。

5 おわりに

AI関連発明の出願は、これまでのソフトウェア特許の書き方のみに沿って準備したのでは不十分なことも多いです。今後も出願増加が予想される中で、審査基準や審査ハンドブック等に基づき、また米国等のAI先進企業の出願ノウハウも参考にしつつ、さらに今後の審決例や判決例にも注意しながら、請求項や明細書を作成していく必要があります。

今後も、本誌においてAI関連発明を継続して取り上げていきます。

【出典】※いずれもウェブサイトより入手可能
1 日本特許庁「AI関連発明の出願状況調査報告書」(2020年7月更新)
2 日本特許庁「AI関連技術に関する事例について(事例の全文)」(2019年1月)

○この記事に関するお問い合わせ先
知財情報戦略室: ipstrategy@soei-patent.co.jp