

特許庁における人工知能技術の活用に関する取組について

—人工知能技術の業務への適用可能性の検討及び電話等の質問対応に関する実証事業を中心に—

JPO's Activities for utilization of AI Technologies

特許庁 総務部総務課調整班課長補佐

麻川 倫広

2008年4月特許庁入庁。特許審査等に従事し、2018年4月より現職。

1 はじめに

産業財産権を取り巻く環境の多様化・複雑化に伴い、特許庁の事務に関しても、料金減免・早期審査等の制度の複雑化や先行調査における調査対象資料の増加等に起因して業務量が増加している。こうした背景を踏まえ、特許庁では、平成28年度から、特許行政の高度化・効率化に資することを目的に、人工知能（AI）技術の特許行政事務への適用可能性の検討を行ってきた。2017（平成29）年4月には、検討の結果を踏まえ、将来的なAI技術の活用を視野に入れ「特許庁における人工知能（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・プランという。）を取りまとめて公表している¹。さらに、平成29年度は、このアクション・プランに沿って、「電話等の質問対応」、「紙出願の電子化」、「特許分類付与」、「先行技術調査」、「先行図形商標の調査」、「不明確な指定商品・役務調査」の6つの業務について実証事業を実施した。

本稿では、まず、アクション・プランの策定にあたって、どのように庁内業務へのAI技術の適用可能性を検討してきたか説明するとともに、平成29年度に実施した実

証事業のうち「電話等の質問対応」に関する事業について具体的な実証内容を紹介する。

なお、本稿中の見解等は、筆者の個人的なものであり、組織の見解等を表すものではない点ご了承ください。

2 AI技術の業務への適用可能性の検討

特許庁の業務へのAI技術の適用可能性を検討するにあたっては、庁内の業務全体を棚卸し、AI技術の適用可否を判断すべき業務の選定を行った。そして、選定された業務について、ソリューション案を立案し、AI技術の技術水準、外部有識者の意見等を踏まえ、導入可能性を検討した。

2.1 業務の棚卸・選定

特許庁内には、出願の受付から審査、登録、審判等に至るまで多岐にわたる業務が存在する。AI技術の業務への適用可能性の検討にあたっては、まず特許庁の業務全体の棚卸しを行った。この結果、892の業務が今回の調査対象の基礎となることを特定した。

次に、各業務について、業務の画一性や業務負荷等の業務分析を行い、AI技術の適用可否を検討すべき業務の選定を行った。具体的には、892の業務のうち、既にシステム化対応済みの業務、あるいは業務プロセスが画一的

1 特許庁における人工知能（AI）技術の活用に向けたアクション・プランの公表について
https://www.jpo.go.jp/torikumi/t_torikumi/ai_action_plan.htm

表 1 AI 技術の適用可否を検討すべき業務として選定した 15 分野の業務一覧

	業務名	対象	業務内容・課題	ソリューション案
1	電話等の質問対応	全課室	担当課室の特定、問合せ対応 (関係課室からの、のべ6業務を統合)	導入可能性あり
2	紙出願の電子化	受付	紙出願された資料を、OCR (光学的文字認識) によりテキスト化	導入可能性あり
3	書類の印影確認	受付・方式	申請書類の印影とマスタ登録された印影の照合	導入可能性あり
4	出願における登録商標の使用の確認	特許	出願書類中で登録商標が使用されている場合に、該当部分を特定	導入可能性あり
5	品質監査 (作成書類の誤記確認)	特許	審査官の作成書類の記載の形式的な瑕疵の確認	導入可能性あり
6	特許分類付与	特許	出願された発明に対し、発明の技術的事項に基づき、適切な分類を付与	一部導入可能性あり
7	先行技術調査	特許	出願された発明に関連する先行技術文献を、データベース等を活用して抽出	一部導入可能性あり
8	発明の内容理解・認定	特許	出願された発明の内容を理解し、特許保護を求めている技術的思想の技術的範囲を認定	現時点では導入困難
9	特許登録可否の判断	特許	出願された発明と過去の発明を対比し、特許を付与すべきかを判断して、判断の具体的理由を示す拒絶理由通知等を作成	現時点では導入困難
10	意匠分類付与	意匠	出願意匠に対し、日本意匠分類・ロカルノ分類を付与	一部導入可能性あり
11	先行意匠調査	意匠	出願された意匠の新規性、創作非容易性の判断に資するための調査を実施	現時点では導入困難
12	意匠登録可否の判断	意匠	先行意匠と出願意匠を対比し、新規性等の要件を満たしているか検討し、登録可否を判断	現時点では導入困難
13	先行図形商標の調査	商標	出願された図形商標と同一又は類似の商標の有無を調査	導入可能性あり
14	不明確な指定商品・役務調査	商標	商標の使用をする商品・役務の記載 (特定) が不明確な出願に対し、どのような商品・役務であるかを調査	一部導入可能性あり
15	商標の審査判断	商標	出願された商標の識別性の有無、不登録事由の有無を検討し、商標登録の可否を判断	現時点では導入困難

で、単純なシステム化で対応可能な業務を除外した。次に、残った業務について、①業務負荷、②業務頻度、③職員の精神的な負荷の3点を評価指標として評価を行った。そして、担当課室へのインタビュー等を経て、指標が高く優先度の高い20業務をAI技術の適用可否を検討すべき業務として選定した。なお、20業務のうち6業務については、各関係課室における電話等の質問対応であり、これらを統合することで15分野の業務とした。

2.2 ソリューション案の検討

業務選定の結果、AI技術の適用可否を検討すべきとされた15分野の業務について、それぞれ業務課題の解決・軽減が期待されるソリューション案を検討するとともに、当該ソリューション案の導入可能性について、AI技術の技術水準から検討を行った。具体的には、担当課室へのインタビュー等を通じて各業務に対する業務課題

を特定し、当該業務課題を解決・軽減しうるソリューション案を機能ごとに立案した。その後、文献調査等を通じ、それぞれのソリューション案について、導入に必要なとなる技術水準を確認することで、導入可能性の可否を検討した。検討に当たっては、大学・独立行政法人等の外部有識者の意見を聴き、その内容の妥当性を確認した。検討の結果、15分野の業務のうち、10業務について、業務の全部または一部について必要なソリューション案の導入可能性有りと判断された。また、残りの5業務については、基礎研究を含め関連技術は存在せず、現時点では導入は困難であるという結論に至った。

2.3 先行的な概念実証

ソリューション案を検討した15分野の業務のうち、導入可能性が高いと判断された「電話等の質問対応」に関して、平成28年度中に先行して概念実証を実施した。

特許庁における、出願人等ユーザからの問い合わせは、年間 30 万件にのぼり、問い合わせの内容は出願から方式審査、実体審査、料金まで多岐にわたる。各業務を担当する部署の職員が各問い合わせに対応しているものの、問い合わせ対応のためにその他の業務が中断されてしまうほか、経験の浅い職員がすぐに回答ができず、ベテラン職員に相談する必要があることなどが課題となっている。

概念実証においては、特定の部署における過去の問い合わせ内容を蓄積し、外部からの問い合わせに対応する職員に回答の候補を提示することで、職員の問い合わせ対応をサポートするモデルを構築し、精度を検証した。この結果、モデルが提示する回答候補の上位 5 位までに正答が含まれる割合が、8 割であることが確認された。

なお、先行的な実証として、このほかに「特許分類付与」に関して、外国特許文献への分類付与に関する調査²も行っているが、本稿では割愛する。

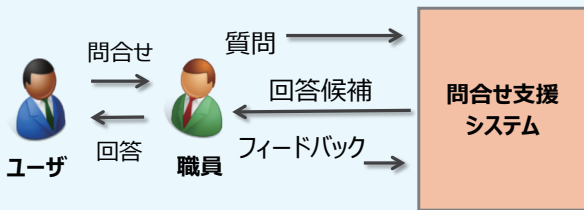


図 1 構築したモデルのイメージ

2.4 アクション・プランの策定・公表

業務の選定、ソリューション案の検討を通じ、庁内業務への AI 技術の適用可能性について一定の結論を得たが、実用化に当たっては AI の学習に必要なデータの収集とそれに基づく実証事業による精度検証等が不可欠である。そこで、AI 技術の業務適用可能性が高いとされた各業務について、技術の適用に向けて必要なプロセスを整理した上で、今後の取組をとりまとめたアクション・プランを策定し、2017（平成 29）年 4 月に公表した。

アクション・プランにおいては、検討したソリューション案について類似の製品・サービスが存在するか基盤となる技術が存在するとされた 10 業務のうち、特に現場のニーズが高い 6 業務すなわち、「電話等の質問対応」、「紙出願の電子化」、「特許分類付与」、「先行技術調査」、「先

2 http://www.japio.or.jp/00yearbook/files/2017book/17_3_01.pdf

行図形商標の調査」、「不明確な指定商品・役務調査」について、必要なデータ蓄積の有無、費用対効果等を確認するため、早期に実証に進めることとした。一方、残り 4 業務について費用対効果の精査を行うこととした。また、現時点では関連技術が存在しない 5 業務についてはソリューション案の見直し、あるいは、当面、AI 技術の進展を注視することとした。

なお、当該アクション・プランは、各事業の今後の取組についておおまかに想定したものであり、実証の進捗状況や関連技術の開発の進展等によって、今後、変更される可能性がある点に留意されたい。

3 電話等の質問対応に関する実証事業

冒頭でも述べたとおり、策定されたアクション・プランに沿って、平成 29 年度に、6 つの業務について実証事業を行った。このうち「電話等の質問対応」に関する実証事業について、ここで紹介する。

2.3 で述べたとおり、「電話等の質問対応」に関しては、職員に回答候補を提示することで問い合わせ対応を支援するモデルについて、平成 28 年度に概念的な検証を行った。これを踏まえ、平成 29 年度の実証事業においては、外部ユーザからの問い合わせにシステムが直接自動で応答するとともに（自動応答）、回答が困難な場合についても、問い合わせ内容を把握したうえで、関係する課室への自動的な課室振り分け（課室転送）を実現するシステムの可能性について検討した。具体的な取組として、システム構築に必要なデータの分析・処理、システムの構築、システム全体あるいはシステムを構成する機能の機械評価を行った。さらに、評価結果を踏まえ、職員による問合せ対応の支援を想定したデモシステムのユーザ評価やシステム案の検討を行った。

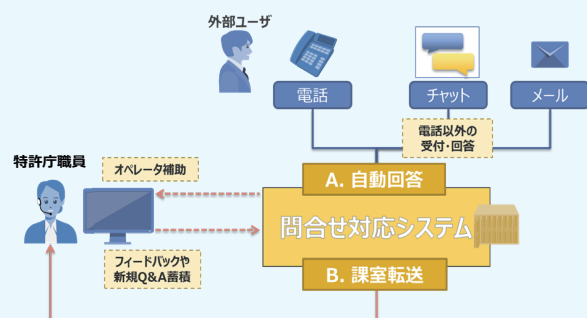


図 2 問合せ対応システムのイメージ

3.1 データの分析・処理

自動応答や課室転送に必要なデータを分析するにあたって、庁内の各課室に寄せられた問合せに関する電話応対メモや電子メールデータを整理し、応対メモ(47,675件)、電子メールデータ(30,397件)を実証事業で活用することとした。これらのデータのうち、本事業の検証に適するデータとして、問合せと回答がペアで記載されている等の条件に合致するものを対象として選定し、応対メモについては、人手によってイメージファイルからテキストデータ化した。そして、システムの自動回答のベースとなる回答知識には、特許庁に蓄積された応対メモ等から計14,000件の問合せと回答の組合せを蓄積した。

また、音声認識等の機能を特許庁業務に適合させるため、特許庁ホームページや一般的なWebサイトを分析し、特許行政事務特有の用語として、合計1,522語を収集した。

そのほか、応対メモの分析結果に基づき、問合せと回答の組合せを自動生成するルールを作成し、特許庁のマニュアル等から539件の組合せを自動生成し、これらについても検証で利用した。このほか、課室転送の実現に向けては、特許庁ホームページにある問合せ先一覧と学習データの有無をもとに分析し、振り分け先となる課室を選定した。

3.2 システムの構築

自動回答あるいは課室転送を実現するためには、合成

音声による受付→ユーザの電話音声のテキストデータ化→質問に対する回答の出力→合成音声による回答又は課室の振り分け、といった処理が必要であり、これらの実現には人工知能技術を用いた複数の機能で検証システムを構築する必要がある。検証システムを構成する機能は、1)合成音声、2)音声認識、3)行動選択、4)言語理解、5)知識探索、6)追加情報要求、7)課室振り分け、8)回答知識獲得である。これらのうち、1)合成音声、2)音声認識、4)知識探索、7)課室振り分けについては、3.1で整備した特許行政事務特有語一覧や蓄積された問合せと回答の組み合わせデータ等を学習データとして用いた。

3.3 機械評価

評価データを用いて、システム全体の機能である自動応答、課室転送、及びシステムを構成する1)~8)の個別機能について機械評価を行った。

評価に用いる問合せと、システムに蓄積されている問合せと回答の組合せの表現の違いから、難易度を4つのレベルで設定し、レベルに合わせて表現を変更した評価データを用意した。

レベル1. 同義語、類似語といった単語レベルでの同一性判定が必要

レベル2. 単語間の関係性を考慮した文の理解が必要

レベル3. 文脈に応じた省略の補完や、追加すべき条件の補完が必要

レベル4. 一般常識による内容の補完が必要

機械評価は計2回実施した。まず、1回目の機械評価

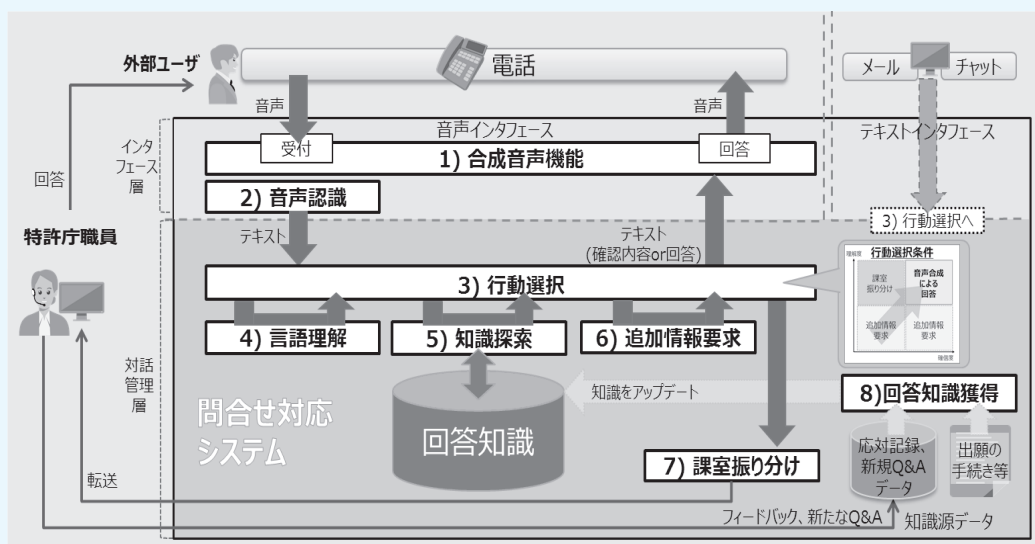


図3 問合せ対応システムの機能構成

を実施し、その結果を踏まえ、自動回答及び課室転送の精度のボトルネックとなる個別機能の対策を図り、2回目の評価を実施した。

評価に際しては、評価データ全体を対象とした場合に加え、検証システムが自動回答可能と判断したケースや課室転送が必要と判断したケースにおいて、それぞれ正しい回答や転送先の課室が提示できたか、という点も評価した。

その結果、評価データ全体を対象とした評価において、自動回答の正答率は1回目の39.9%から2回目は46.8%、課室転送の正答率は1回目の27.8%から2回目は36.4%と、2度の検証で精度の向上が認められた。また、自動回答について、システムが自動回答可能と判断したケースにおける正答率は、1回目の57.8%から2回目は79.9%と、大幅に精度が向上した。課室転送については、課室に関する学習データが多く存在した一部の課室で70%を超える精度が見られたが、総じて精度としては低い結果となった。なお、ここでの正答率は、システムが提示する1位の回答候補に正しい回答があがっていることを指す。

精度の向上した自動回答について、検証観点で示した4つのレベル別の正答率を見ると、レベル1及び2の2回目の評価では、システムが自動回答可能と判断したケースでいずれも約95%であるのに対し、問合せの条件が部分的に欠落したレベル3や条件の欠落・表現が抽象化されたレベル4では、2回目の評価においても、ほぼ正しい回答を提示できていないことが判明した。

検証システムは人工知能技術を用いた複数機能で構成される。このため、システム全体の精度のボトルネックとなる機能を特定するために、機能別の検証項目を設定し、評価した。1回目の機械評価で、「1)音声合成」「2)音声認識」は高い精度であったものの、その他の機能がボトルネックとなっていることが判明した。

ボトルネックとなった機能については、対策の難易度及び想定効果から機械評価2回目に向けた対策実施の優先度を設定し、優先度の高いものから対策を施した。その結果、応答の中核となる「4)知識探索」において改善効果を確認できたが、その他の機能では大きな改善効果は見られなかった。

また、改善効果の高い知識探索機能について、N位

正解率のNを変化させた場合、N=5（上位5位までに正解を出力できた割合）では、1回目の66.1%から86.3%に向上することを確認した。

音声による自動応答では、複数の回答候補を提示することが困難であるため、N=1における高い精度が求められる。一方、職員による問合せ対応を支援するシステムやチャットによる応対システムでは、上位の複数候補から利用者が最適な回答を選択する方式も一般的に採用されており、当該評価結果から音声による自動応答以外のシステム適用可能性もあることを確認した。

3.4 デモシステムのユーザ評価

機械評価の結果を踏まえ、職員による問合せ対応の支援を想定したデモシステムを用意し、問合せ対応業務に従事する特許庁職員を対象としたユーザ評価を行った。

デモシステムは、外部ユーザからの問合せを音声認識機能によってテキスト化し、テキスト化した問合せ内容を基に、問合せと回答の組み合わせが蓄積されたデータベースから、知識探索機能によって類似する組み合わせを抽出して、職員が利用する画面に回答候補として表示する構成とした。また、デモシステムの利用に際しては、職員にヘッドセットを装着した応対を試行した。評価はデモシステムを試行した後に、視認性、簡索性、レスポンス性、操作性、全体満足度の観点についてアンケート形式で実施した。

レスポンス性について、即応が求められる電話対応においては、より一層のレスポンス性能が求められることが判明した。今回のデモシステムは、問合せ内容をリアルタイムで把握して、都度、蓄積されたデータベースを参照する仕組みとしたことから、一定の会話のまとまりを捉えるために、一般的に会話の中で生じる切れ目を考慮した待ち時間を数秒セットした。システム導入に向けては、実業務での試行を通じて、このような待ち時間や情報探索性能などの要件を明確化する必要がある。操作性については、回答の根拠となるマニュアル等を提示する機能の他、問合せ内容と回答候補で合致する箇所をハイライトする等の視認性を向上するための機能などが求められた。

また、全体満足度については、ユーザ検証時の音声認識精度の低さへの不満が多くあげられた。機械評価時は

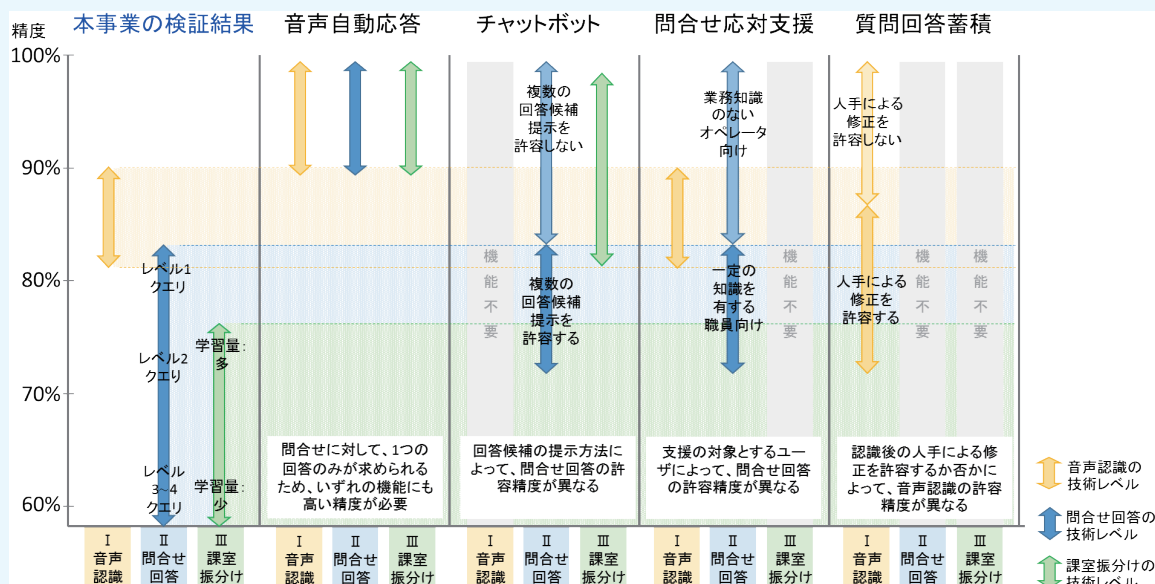


図4 導入システム案の各構成要素に求められる精度

高い認識精度であったことに鑑み、ユーザ評価で取得した音声データを分析したところ、半数のデータで音割れやノイズの発生が確認された。このため、再現検証を行った結果、これらの多くはヘッドセットの装着漏れ、マイク端子との距離、マイク端子への接触等による影響が大きいことが判明した。実用に当たって、同様にヘッドセットなどの新たな機材を導入する際は、これら機材の利用方法の教育にも考慮が必要と考えられる。

3.5 システム案の検討

機械評価及びユーザ評価の結果をもとに導入システム案を検討した。本事業で目指す問合せ対応システムは、音声認識、問合せ回答、課室振分けの3つの要素から構成される。これらの要素を組み合わせることで、音声自動応答システム、チャットボットと呼ばれるテキスト自動応答システム、問合せ対応支援システム及び質問回答蓄積システムの4パターンの実現可能性が考えられる。

これらの導入システム案は、それぞれ実現方法によって、各構成要素に求められる精度も異なる。本事業における検証結果と、導入システム案に求められる精度を比較した結果、チャットボットシステム、問合せ対応支援システム、質問回答蓄積システムは、実現可能性ありと判定した。

3.6 考察

検証の結果、問合せのレベルに応じて自動回答の実現可能性に違いがあることが判明した。同義語や単語の入

れ替えといった違いを持つ問合せに対しては高い精度で対応可能な一方、内容を特定する条件が不足し、問い返しを必要とする問合せに対しては、現時点で自動回答は困難であることが判明した。

このため、システムの段階的な導入によって、職員の負担軽減や外部ユーザの利便性向上を図るとともに、対応可能な問合せの範囲拡大や精度向上に向けた学習データの蓄積を進めることが重要と考えられる。また、段階的な導入に当たっては、ターゲットとする利用者（職員及び外部ユーザ）を定め、その要望を取り込むことが必要であり、特に職員に向けては導入段階で現場ニーズを洗い出すだけでなく、小規模での実運用からはじめ、継続的な改善を回す仕組みが求められる。

さらに、本事業の検証では、対応メモ等から作成した蓄積データを活用する上で、問合せ内容や回答内容の記載が断片的であるなど、自動回答の利用を想定した場合におけるデータの質に関する課題が確認された。今後、システム導入に向けて、これらの膨大なデータの整備を進める上で、人手による確認作業は不可欠であり、それを支援する技術の検討も必要と考えられる。

4 今後の予定

「電話等の質問対応」については、平成29年度の実証事業における成果、課題を踏まえ、平成30年度も引き続き実証事業を実施する予定である。