

生成 AI を用いた知財金融向け特許情報活用支援技術の検討 Examination of patent information utilization support technology for intellectual property finance using generative AI

谷川 桂子[†] 郡浦 宏明[†] 細矢 淳[†]
Keiko Tanigawa Hiroaki Konoura Jun Hosoya

1. はじめに

金融機関は、企業の経営課題の把握、方策の策定、適切なアドバイスやファイナンスの実施を通して、地域経済発展へ貢献することを求められている[1]。金融庁も金融機関が地域企業への金融仲介機能を発揮できるよう様々な施策を取っている。例えば、金融機関の融資は企業の財務情報や担保情報などに依存する傾向があるため、これに寄らない事業性評価に基づく融資がなされているかを検証・公表し、金融機関の現場職員のノウハウを共有することで、企業支援能力向上の取り組みを後押ししている。

事業性評価のために活用できる情報として、知的財産(以下、知財)が挙げられている。特許庁も知財をもとに取り引先企業の強みを活かす施策「知財金融」を推奨しており、知財金融ポータルにてノウハウの共有やコンサルタントを提供している[2]。

以上のように、金融機関に対して知財の積極活用が求められているが、専門家でも読み解きの難しい特許文書を現場の銀行職員が理解するにはハードルが高い。本研究では、この点に鑑み、金融機関が知財のうち特許を活用する際の読み解き支援技術として、生成 AI を活用する検討を行った。本稿では、地域企業の成長に向けた提案活動の一連のプロセスを仮定し、各プロセスにおける作業、困り事、それを解決する支援技術をユースケースとして整理した結果と、生成 AI を用いた特許情報活用支援プロトタイプ概要について報告する。

2. 特許情報活用支援技術の検討

2.1 中小企業向け新領域進出支援のプロセス

地方銀行には地域の取引先企業の強みを把握した上で成長に向けた提案が求められている。企業技術の新領域候補の提案は、(1)取引先企業の技術把握、(2)競合企業の技術との比較、(3)技術動向の把握、(4)新領域候補の提案の4つのプロセスを経るものと想定した。

上記プロセスにおいて多くの時間が必要な作業の一つが情報収集である。企業情報を得る手段としてインターネットを通じた企業の Web ページからの取得があるが、地方銀行の取引先企業を中心とする中小企業が必ずしも適宜情報を発信しているとは限らないため、必ず取得可能なオープン情報が必要である。知財は事業の競争力の源泉であり、強み技術を文書化した有用なオープン情報であるため、企業技術把握のデータソースとして特許を活用する。特許公報を要約することで(1)(2)取引先/競合企業技術の把握や(4)新領域候補の推定に活用できる。また、(4)新領域候補の推定には、特許公報の他に技術分野の定義が必要である。これには総務省が公開している日本標準産業分類[3]を用いる。(3)技術動向の把握については Web 検索により取得可能で

はあるが、データソースの信用度を担保できないため、本検討では対象外とする。

2.2 特許情報活用における問題

技術把握、産業分類推定共に、特許活用は特許公報の読み解きから始まる。特許公報は技術文書であるため、専門技術が記載されている。その上、権利を広く取るために抽象的な表現が多く一文が長くなりがちで、専門家にとっても難解である。背景知識のない金融機関の現場職員にとっては更にハードルが高くなる(図1)。特許公報を知見のない人にも分かりやすい文章にかみ砕いた表現にする必要があり、これに生成 AI を活用する。

2.3 特許情報活用支援機能

2.3.1 アーキテクチャ

特許情報の読解に生成 AI を活用する。生成 AI はオープンデータの学習済みモデルであるが、特許公報や産業分類定義などは学習の対象外となっている。このため、外部データを生成 AI の回答生成に活用する技術 RAG(Retrieval Augmented Generation)を用いて、特許公報と産業分類定義を取り込む。図2にアーキテクチャを示す。

特許情報活用支援機能は、外部データを取り込むための RAG Preprocessing, RAG Postprocessing と、ユーザが入力する問合せに対して生成 AI を活用するための Prompt Engineering より構成する。



図1 特許活用における問題

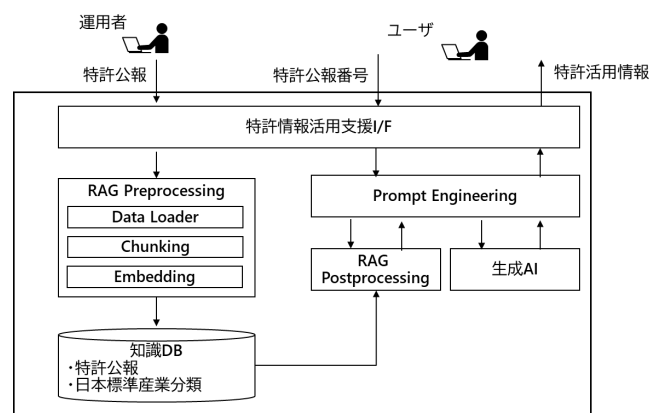


図2 特許情報活用支援機能のアーキテクチャ

[†] 株式会社日立製作所, Hitachi, Ltd.

以上を考慮し、特許公報全文を対象に、要約の観点(例えば分野、用途、特徴技術、製品例など)をプロンプトに指示し、技術のエッセンスを抽出することを目指した。注意点は、特許公報には専門用語が多用されており、当然ながら生成 AI の回答文にも専門用語が含まれるため、それらを解釈させる指示も必要となることである。

2.3.3 特許公報比較

取引先企業の技術とその競合企業の技術の比較は、新領域分野の候補を検討する上で有用である。生成 AI に特許公報を比較させる手段は2つある。

①要約を比較

2.3.2 で要約された技術を比較する。

②特許分類コードを比較

特許公報には、特許庁審査官など専門家が付与する分類コードが記述されている。分類コードはその特許の技術的な特徴を表すものであるため、これを技術の比較に活用できる。

②は特許公報から分類コードを抽出し、その定義情報を比較する処理で済むため、①よりも更に処理時間を低減できる。比較対象の特許数が多い場合は手段②を実施するなど、サービス要件によって使い分けが可能である。

2.3.4 産業分類推定

産業分類推定は、取引先企業の技術が活用可能な分野の把握に活用する。2.3.1 で生成した要約で示される技術が、産業分類定義のどの細分類項目に該当するのかが推定する。本検討で用いた日本標準産業分類の細分類は 1,460 項目であり、各項目に属する複数の産業のうち特許技術と関連のある細分類を抽出させるための指示が必要である。更に細分類項目が多いため、回答のばらつきや揺らぎが起りやすいことが懸念される。推定された産業分類が問合せ毎に異なると、回答の信頼性に疑問が残る上、確認範囲が広がるり、余分な工数が発生する。そこで、一つの要約技術に対して産業分類推定を複数回行い多数決を取ることにした。まず、各産業分類推定において、生成 AI が推定した複数の細分類項目をスコア順に上位から最大 3 件応答させる。これを複数回、例えば 5 回繰り返すことで、ランクの入れ替えはあれど、ある程度推定結果のばらつき・揺らぎが収束することが期待できる。ランク毎に重みを付けて 5 回分の総和によって確度を担保する。

(S) Int. Cl.	F I	テーマコード (参考)
B 2 2 C 9/04 (2006.01)	B 2 2 C 9/04 R	4E093
B 2 2 C 9/22 (2006.01)	B 2 2 C 9/04 N	
B 2 2 D 25/02 (2006.01)	B 2 2 C 9/22 Z	
	B 2 2 D 25/02 A	

記号	タイトル
B	処理操作・運輸
B22	鍛造・熱処理
B22C	鍛造用鋼型造型 (耐火物の造型一般 B 2 2 B)
<注>	このサブクラスは、以下のものを包含する: 一 金属の鍛造用鋼型またはその他の耐火物鋼型の製造; 一 鋼型用材料の選択と調整; 一 必要な原型、製造方法、構成、付属装置または工具
<索引>	原型、その製作 7/00, 3/00 鋼型、中子、一般的な造型方法; 鋼型と中子の組立て 9/00; 1/00, 3/00 造型機械、造型機械を伴った方法 11/00~19/00 造型装置 2/00 工具または他の装置 5/00, 21/00, 23/00 耐火物鋼型の組成物または中子材料; その粒子構成 (耐火物造型一般 C 0 4 B 3 5/00); 鋼型成形または製作上の化学的または物理的特徴
B22C 1/00	耐火物鋼型の組成物または中子材料; その粒子構成 (耐火物造型一般 C 0 4 B 3 5/00); 鋼型成形または製作上の化学的または物理的特徴
B22C 1/02	特別な目的の添加物を特徴とするもの、例、指示薬、腐蝕用添加剤
B22C 1/04	鍛造物を保護するもの、例、脱炭素に対するもの
B22C 1/06	短縮な酸化性の金属を排除するためのもの
B22C 1/08	鋼型の腐蝕を避けるためのもの、例、インベストメント鍛造用のもの

図 4 特許分類コードと分類コードの定義の例

3. 検証

本研究では、銀行職員が(1)技術概要把握、(2)産業分類推定を行う際の手動作業に対する生成 AI 活用時の工数削減効果と、生成 AI の回答の分かりやすさを検証する。

3.1 検証条件

3.1.1 検証環境

(a)AI モデル

AI モデルは特許公報や産業分類定義が日本語であるため、多言語対応のベンチマーク結果が良い下記のモデルとした [4-6]。Temperature は生成 AI の回答の揺らぎを抑える手法の一つである。0.0~1.0 の値を取り、1.0 に近いほど多様な回答を得られる。本検討では日本語としてできるだけ違和感がなく、また一貫性のある回答を得るため、以下の設定とした。

- AI モデル
 - LLM(Large Language Model) : GPT-4o
 - Embedding : text-embedding-3-large
- パラメータ
 - Temperature : 0.0

(b)プロンプト例

(1)技術概要把握

あなたは優秀な弁理士です。与えられた特許公報の特徴を、背景知識のない人が理解できる単語で説明してください。特許公報の内容を、対象分野、この特許が提供する技術、特徴、効果を、それぞれ行で、日本語で回答してください。技術用語には説明を加えてください。

(2)産業分類推定

あなたは優秀な弁理士です。あなたの仕事は、特許技術を活用できる産業分類を選択することです。以下の特許の内容と特許分類に基づき、条件に従った産業分類を推定し、細分類の番号とその項目をスコア順に最大3つ選択し、出力形式に従って教えてください。

#特許公報の内容
#summary
#制約
・管理事務を行う本社や事業所は除外する。
・特許の特徴に該当しない産業分類は除外する。
#出力形式
(1)aaaa_細分類項目;理由
(2)bbbb_細分類項目;理由
(3)cccc_細分類項目;理由

3.1.2 作業工程見積もり

人手による(1)技術概要把握と比較、(2)産業分類推定にかかる作業工数を以下に見積もった(図 5)。

日本人の文書の平均読み取り速度は 400~800 文字/分と言われている[7]。読み取り速度を 500 文字/分、特許公報平均 20 ページ、1500 文字/ページとすると、特許公報読み取りには 1 公報あたり約 1 時間要する。一般的に知見のない技術文書の読み取りは小説等の文書を読む以上の時間がかかるものと想定できるため、3 倍に設定した。1 公報の技術把握に要する時間は約 3.5 時間、競合企業技術との比較までに要する時間は約 8.5 時間、産業分類推定には約 4 時間とした。

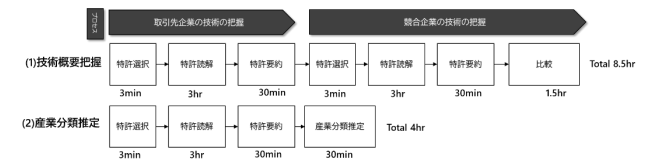


図 5 人的作業工数見積もり

3.2 検証結果

3.2.1 工数削減効果

生成 AI 活用による工数削減効果を図 6 に示す。

(1) 技術概要把握と比較

特許情報支援機能を利用することで、ユーザの作業は、比較させる 2 つの特許公報を選択して公報番号を入力するだけでよく、合計 8.5 時間かかっていた作業を約 17.5 分で行うことができる。特許公報の RAG 化(ベクトル化)に時間がかかるが埋め込みモデルの進化や、対象特許公報番号を予めリスト化・バッチ処理することで処理時間の更なる短縮を期待できる。

(2) 産業分類推定

(1)同様、ユーザの作業は特許公報を選択して特許公報番号を入力するだけでよく、4 時間かかっていた作業を約 10 分で行える。本検証では実験環境であること、産業分類の回答揺らぎ吸収のための 5 回分の処理時間を含んでいる等の理由でリアルタイム利用するには時間がかかっている。生成 AI モデルや、回答の一貫性を確保する技術等の生成 AI 周辺の連携技術の進化による処理時間の短縮を期待できる。

3.2.2 生成 AI 回答の分かりやすさ

(1) 特許要約

初心者向け特許要約の提示例を示す。なるべく平易な単語が使われた文章となっており、有識者より、分野や製品

【分野】	リチウムイオン電池などの電池に使われる新しい材料の開発。
【ターゲット】	電池の性能を向上させるための、シリコンと炭素を組み合わせた新しい材料。
【特徴】	シリコンで覆われた炭素の小さな粒(ナノサイズのカプセル)を使い、電池の中での膨張や劣化を抑える工夫がされている。
【効果】	電池の容量が増え、長持ちするようになり、軽くて小さい電池が作れるようになる。

イメージをつかむには十分であるが、知見レベルによって分量や難易度を調整できると良いとの評価を得ている。

(2) 特許公報比較

特許分類コードの定義から 2 つの特許公報を比較した例を示す。分類コード定義は技術的用語のみの場合が多いため、補足説明を含んだ文章になっている。

【類似点】	金属を鋳造する方法の発明であり、特殊な形状の鋳物を使用することに特徴があります。また、一度使うと消える原型や溶ける素材を使用する点でも共通しています。
【相違点】	特許1は、網状の消失模型という消える原型を鋳物の中に埋め込むことで特殊な形状の鋳物を作ります。一方、特許2は、溶ける素材(水溶性中子)を使って鋳型を作り、その中に溶融金属を流し込むことで特定の形状の金属製品を作り、その素材を活かすことで金属製品の中に規則的な空洞(開放孔)を作り出します。

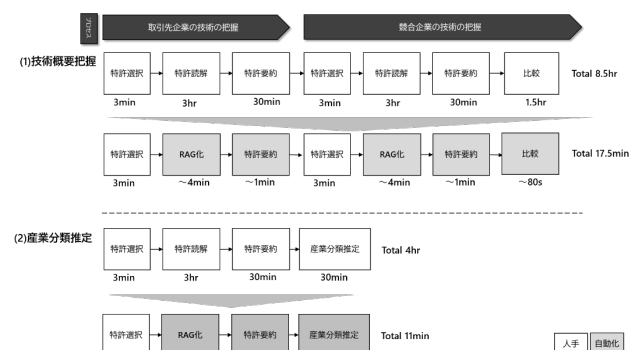


図 6 生成 AI 活用による工数削減効果

(3) 産業分類推定

2.3.4 で想定した通り、問合せ毎にランクや細分類項目が異なるケースがあった。これを文章で記述するとユーザに提示する情報が多くなり、一目で分かる可視化手段が必要である。図 7 に一例を示す。(1)で示した特許技術はリチウムイオン電池などの新素材に関するものなので、蓄電池製造業が高ランクで選定されていることが分かる。また、生成 AI はなぜこの細分類を選択したのかという、選択理由を併せて提示することも可能である。有識者からは、関連産業の調査のきっかけの情報として利用可能であり、有用であるとの評価を得た。

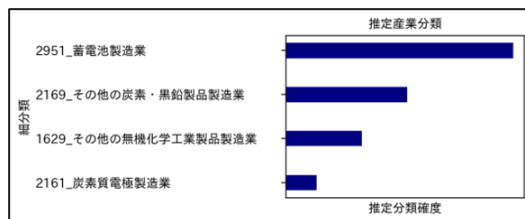


図 7 産業分類推定

4. おわりに

本稿では、銀行職員の地域企業への支援活動の工程を仮定し、取引先企業や競合企業の技術把握及び展開先産業分野の候補の推定に生成 AI を活用した特許情報活用支援機能の開発と、その効果を検証した。初心者に向けた特許情報の分かりやすい説明文章を作成させる試みや生成 AI の課題であるハルシネーションへの対策について述べた。読解のハードルの高い特許文書を容易に把握できると共に、従来知見のない分野や技術文書の読解に要する工数を劇的に削減できる。

今後は知見レベルに合わせた分量や難易度の文章の生成や図表などテキスト以外の情報を取り込み、より分かりやすい情報の提供を目指す。

参考文献

- [1] 金融庁, “事業全体を対象とする担保制度の検討,” https://www.fsa.go.jp/policy/jigyuu_tanpo/index.html
- [2] 特許庁, “知財金融ポータルサイト,” <https://chizai-kinyu.go.jp>
- [3] 総務省, “日本標準産業分類,” https://www.soumu.go.jp/toukei_toukatsu/index/seido/sangyo/H25index.htm
- [4] OpenAI, “Hello GPT-4o,” <https://openai.com/index/hello-gpt-4o/>
- [5] Microsoft, “Azure OpenAI Service models,” <https://learn.microsoft.com/en-us/azure/ai-services/openai/concepts/models?tabs=global-standard%2Cstandard-chat-completions>
- [6] Sakib Shahriar et al., “Putting GPT-4o to the Sword: A Comprehensive Evaluation of Language, Vision, Speech, and Multimodal Proficiency,” <https://arxiv.org/pdf/2407.09519>
- [7] SP 速読学院, “あなたの読書スピードと、日本人平均値の関係,” <https://www.pc-sokudoku.co.jp/kouza/kougo/kyohon29.html>