

機械学習を用いた効率的な特許調査方法

—大規模言語モデルによる特許調査の効率化—

Effective patent search methods using Machine Learning



花王株式会社 研究開発部門 知的財産部/アジア特許情報研究会

安藤 俊幸

1985年現花王株式会社入社、研究開発に従事
 1999年研究所の特許調査担当（新規プロジェクト）、2009年知的財産部
 2011年よりアジア特許情報研究会所属
 2020年 特許情報普及活動功労者表彰 日本特許情報機構理事長賞「技術研究功労者」受賞
 情報科学技術協会、人工知能学会、データサイエンティスト協会 各会員

✉ ando.t@kao.com

1 はじめに

特許調査は、競合他社の特許や技術トレンドの把握、自社の知財ポートフォリオの評価など、企業にとって重要な活動であるが、時間とリソースを必要とする課題も存在する。これらの特許調査の課題を緩和するような複数のAI利用特許調査ツールが商用利用されている。また最近、世界的にもChatGPTをはじめとした対話型AI（生成系AI）がポジティブ・ネガティブ両面で大変注目を集めている^{1,2)}。また大学等の教育機関へも大きな影響を与えている³⁾。本稿では、知財分野におけるAI活用による特許調査の効率化について、AI利用特許調査システムや大規模言語モデルを利用したシステム（ChatGPT、BingAI、BERT、Google Bard）を実際に使用して可能な限り検証を行いエンドユーザーの立場としてその可能性を検討する。

2 大規模言語モデルの概要

大規模言語モデル（LLM：Large Language Model）は、自然言語処理（NLP：Natural Language Processing）の分野で注目されている高度な機械学習モデルのことである。これらのモデルは、大量のテキストデータを学習して言語の構造や意味を理解し、自然言語のタスクに対して高度な予測や生成を行うことができる。

大規模言語モデルの代表的な例としては、OpenAIのGPT（Generative Pre-trained Transformer）

やGoogleのBERT（Bidirectional Encoder Representations from Transformers）がある。これらのモデルは、Transformerと呼ばれる深層学習モデルのアーキテクチャを使用しており、自然言語の文脈を理解するために文内の単語やフレーズの関係性を学習する。

大規模言語モデルは、事前学習（Pre-training）と微調整（Fine-tuning）という2つのステップから構成される。事前学習では、大規模なテキストコーパスを用いてモデルが一般的な言語パターンや文脈を学習する。その後、特定のタスクに対して微調整を行い、具体的なタスクに適したモデルを作り上げる。

大規模言語モデルは、自然言語の理解、文章生成、要約、機械翻訳など、さまざまなNLPタスクにおいて優れた性能を発揮する。特に、大規模なコーパスから学習することで、一般的な言語の知識を獲得し、より多様な文脈での文理解や生成が可能となる。

大規模言語モデルは、その豊富な表現力と柔軟性から注目を浴びており、情報検索、対話システム、知識獲得、特許調査など、様々な領域で効率化や進歩をもたらしている。

実際に大規模言語モデル利用したツールを使用してみるとその能力に驚くことも多いが、問題点も多いのが実情である。使いこなすにはツールの特徴の把握と、出力結果の検証は必須である。対話型AIは使い始める敷居は低く簡単に始められる。クエリに対して自然文で回答が返され分かりやすいが、回答内容には間違いが含まれることも多くファクトチェックは必須である。専門的な

内容の検証には手間がかかるが、詳しくない分野のバックグラウンドの把握や予備的な調査には様々なリスクに留意してうまく使うと有用である。

3 大規模言語モデルの具体例

大規模言語モデルとは、多数のパラメータ（数千万から数十億）を持つ人工ニューラルネットワークで構成されるコンピュータ言語モデルで、膨大なラベルなしテキストを使用して自己教師あり学習または半教師あり学習によって訓練が行われる。

大規模言語モデルの具体例として、まず生成系 AI の OpenAI の ChatGPT⁴⁾、Microsoft BingAI⁵⁾、Google Bard⁶⁾ を選択して検討した。BERT に関しては別途検討した。ChatGPT、Microsoft BingAI、Google Bard の特徴をそれぞれに訊ねた結果を表形式にまとめて図 1 から図 3 に示す。

ChatGPT、Microsoft Bing AI、Google Bard の特徴

プロンプト：
ChatGPT、Microsoft BingAI、Google Bard の特徴を表形式にまとめてください。
ChatGPT 3.5 の回答
以下は ChatGPT、Microsoft Bing AI、Google Bard の特徴を表形式にまとめたものです。

特徴	ChatGPT	Microsoft Bing AI	Google Bard
開発元	OpenAI	Microsoft	Google
目的	自然言語処理、対話生成	検索エンジン、自然言語処理、質問応答	音楽、詩の生成
学習方法	事前学習と転移学習	様々なデータソースを活用して学習	様々なデータソースを活用して学習
用途	テキスト生成、対話システム、AIアシスタント	検索結果の提供、質問応答、自然言語処理	音楽生成、詩の生成
機能	自然言語処理、テキスト生成、対話生成	検索エンジン、自然言語処理、質問応答	音楽生成、詩の生成
応用	チャットボット、文章作成支援、AIアシスタント	ウェブ検索、情報検索、質問応答、AIアシスタント	音楽作曲、詩の創作

上記の表は、ChatGPT、Microsoft Bing AI、Google Bard の特徴を要約しています。それぞれのモデルは異なる目的と機能を持っており、開発元によって異なるアプローチが採用されています。

図 1 OpenAI の ChatGPT の回答

ChatGPT の回答を Excel にコピーして黄色は筆者が回答に違和感を覚えて設定した。ChatGPT の大規模言語モデルは 2021 年 9 月までの情報に基づいているためと考えられる。

BingAI の回答では言語モデルの大きさに関して、大規模言語モデルのパラメータ数の意味とパラメータ数自体に疑問を感じたので ChatGPT に以下のような質問をした。筆者「大規模言語モデルのパラメータ数の意味を教えてください。」

ChatGPT の回答

「大規模言語モデルのパラメータ数は、そのモデルが学習中に調整可能な重みとバイアスの数を表します。これらのパラメータは、モデルがデータからパターンや関

ChatGPT、Microsoft Bing AI、Google Bard の特徴

プロンプト：
ChatGPT、Microsoft BingAI、Google Bard の特徴を表形式にまとめてください。

Microsoft BingAI の回答
ChatGPT、Microsoft BingAI、Google Bard の特徴を表形式にまとめました。以下の表をご参照ください。

ChatGPT、Microsoft BingAI、Google Bard の特徴を表形式にまとめました。以下の表をご参照ください。

特徴	ChatGPT	Microsoft BingAI	Google Bard
開発元	OpenAI	Microsoft	Google
言語モデルの大きさ	13億パラメーター	2.5億パラメーター	4.6兆パラメーター
リリース年月日	2020年6月	2021年11月	2023年3月
日本語対応状況	対応済み	対応済み	対応済み

以上です。ご参考になれば幸いです。

図 2 Microsoft BingAI の回答

係性を学習し、入力に対する適切な出力を生成するために使用されます。

パラメータ数が多いほど、モデルはより多くの情報を記憶し、より複雑なパターンや関係性を学習することができます。大規模なパラメータ数は、言語モデルの表現力や生成能力を向上させる可能性があります。

しかし、大規模なパラメータ数はコストやリソースの面でも課題を抱えます。モデルの学習や推論には膨大な計算資源が必要となり、モデルのサイズや処理時間が増加します。また、大規模なパラメータ数を持つモデルはデプロイメントやメモリ制約の面でも挑戦をもたらすことがあります。

したがって、大規模言語モデルのパラメータ数は、そのモデルの表現力や計算コスト、リソース制約などを考慮しながらバランスを取る必要があります。」

この説明により、パラメータ数の意味は概ね理解できた。

具体的なパラメータ数自体は間違った回答や矛盾した回答で ChatGPT からは正解は得られなかった。

筆者「GPT3 のパラメータ数を教えてください」

ChatGPT 「GPT-3 は、1.75 兆 (1,750 億) 個のパラメータを持つ大規模なモデルです。」

「1.75 兆 (1,750 億) 兆と億の関係が間違っている。

筆者「GPT3.5 のパラメータ数を教えてください」

ChatGPT 「GPT-3.5 は、1750 億個のパラメータを持つモデルです。」。GPT3 と GPT-3.5 の違いも分かっていない。

筆者「GPT4 のパラメータ数を教えてください」

ChatGPT 「申し訳ありませんが、私の知識は 2021 年 9 月までの情報に基づいており、GPT-4 の具体的なパラメータ数については提供できません。GPT-4 の詳細や具体的なパラメータ数については、公式発表や関連



情報を参照することをおすすめします。GPT シリーズは進化を遂げており、将来のバージョンではより大規模なパラメータを持つ可能性があります。」

回答内容のファクトチェックは必須だが内容によってはかなり困難な場合もある。ちなみにGPT4のパラメータ数は非公開である。

ChatGPT、Microsoft Bing AI、Google Bardの特徴

プロンプト：
ChatGPT、Microsoft BingAI、Google Bardの特徴を表形式にまとめてください。

Google Bardの回答
はい、日本語で答えます。

機能	ChatGPT	Microsoft BingAI	Google Bard
開発者	OpenAI	Microsoft	Google AI
トレーニングデータ	テキストとコード	テキストとコード	テキストとコード、Google 検索結果
パラメータ数	175B	137B	1.56T
アーキテクチャ	Transformer	Transformer	Pathways
能力	テキスト生成、言語翻訳、質問応答、コード生成、クリエイティブコンテンツの作成	テキスト生成、言語翻訳、質問応答、コード生成、クリエイティブコンテンツの作成	テキスト生成、言語翻訳、質問応答、コード生成、クリエイティブコンテンツの作成、Google 検索結果に関連する情報を生成、ユーザーの質問に回答
利用状況	開発者向け	開発者向け	一般公開

hatGPT、Microsoft BingAI、Google Bard はすべて、大規模な言語モデル (LLM) と呼ばれる機械学習モデルです。LLM は、テキストとコードの膨大なデータセットでトレーニングされており、テキストを生成し、言語を翻訳し、質問に答えたり、コードを生成し、クリエイティブコンテンツを作成したりすることができます。

ChatGPT と Microsoft BingAI はどちらも、OpenAI と Microsoft によって開発された LLM です。ChatGPT は 175B のパラメータ数を持ち、Microsoft BingAI は 137B のパラメータ数を持ちます。どちらのモデルも Transformer アーキテクチャを使用しています。

Google Bard は、Google AI によって開発された LLM です。Google Bard は 1.56T のパラメータ数を持ち、Pathways アーキテクチャを使用しています。Google Bard は、テキスト生成、言語翻訳、質問応答、コード生成、クリエイティブコンテンツの作成、Google 検索結果に関連する情報を生成、ユーザーの質問に回答することができます。

ChatGPT、Microsoft BingAI、Google Bard はすべて、LLM の分野で大きな進歩を遂げています。これらのモデルはさまざまなタスクで、人間のレベルのパフォーマンスを発揮することができます。今後も、LLM の研究開発は進んでいくと思われます。

図 3 Google Bard の回答

黄色のセルは気にかかった箇所である。ChatGPT と同様に Google Bard にも質問した。回答を以下に記す。

「GPT-3.5 のパラメータ数は、3,550 億個です。GPT-3.0 は 1,750 億個のパラメータ数を持っていたので、GPT-3.5 のパラメータ数は 2 倍になっています。」「まだ発表されていないので、GPT-4 のパラメータ数はわかりません。」

4 大規模言語モデルの有効性検証

特許調査における大規模言語モデルの有効性検証は、まず生成系 AI の OpenAI の ChatGPT、Microsoft BingAI、Google Bard で行った。検証対象として特許調査のプロセスと正解公報が詳細に解説されている特許検索競技大会 2021 年過去問⁷⁾の化学・医薬分野の問題(徐放性マイクロニードル) 図 4 を使用した。検索結果の性能の指標として正解公報 9 件(図 5) を使用して、正解率、精度、再現率、正解公報のランキング順位を求めた。

図 6 に ChatGPT(GPT3.5) の 2023 年 4 月時点の回答を示す。一見するとそれらしき回答を示しているように見えるが商用の特許データベース CyberPatent Desk を使用して出願番号照会を行うと例示されている

特許検索競技大会2021年過去問の化学・医薬分野の問題 徐放性マイクロニードル

タンパク質等の薬理活性物質を経皮吸収させる技術としてマイクロニードルシートが注目されている。A社ではその改良技術として、徐放性マイクロニードルシートの開発を進めてきた。その結果、開発内容を特許出願することが決定したため、出願に先立ち先行技術調査を実施することになった。以下、A社による特許出願の発明の名称、特許請求の範囲の案を示す。

【発明の名称】徐放性マイクロニードル
【特許請求の範囲】
【請求項1】
水溶性固形成分を基材とする経皮投与用マイクロニードルであって、前記マイクロニードルは内部に微粒子を含み、前記微粒子は乳酸ポリマー、グリコール酸ポリマー及び乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーと薬理活性物質を含有する経皮投与用マイクロニードル。

請求項1の構成要件の分節

- 水溶性固形成分を基材とする(経皮投与用マイクロニードルであって、)
 - 前記マイクロニードルは内部に微粒子を含み、
 - 前記微粒子は乳酸ポリマー、グリコール酸ポリマー及び乳酸-グリコール酸コポリマーから選択される少なくとも1つのポリマーと
 - 薬理活性物質を含有する
- e 経皮投与用マイクロニードル。

図 4 特許検索競技大会 2021 年の過去問

特許検索競技大会2021年過去問の化学・医薬分野の正解

正解公報例

No.	公開・公表番号	発明の名称	出願人(公開時)
1	特表2017-528434	レチノールまたはレチノール誘導体を含むマイクロニードル	エルジーハウスホールディングヘルスケアリミテッド
2	特表2018-510883	タンパク質またはペプチド伝達用の溶解性マイクロニードル	エルジーハウスホールディングヘルスケアリミテッド
3	特表2018-510886	難溶性薬物伝達用の溶解性マイクロニードル	エルジーハウスホールディングヘルスケアリミテッド
4	特開2016-087474	マイクロニードル及びマイクロニードルパッチ	セラジェクト、インコーポレイテッド
5	特表2009-507573	薬物粒子および/または薬物を吸着した粒子を含む、固溶体穿孔装置	セラジェクト、インコーポレイテッド
6	特開2012-110726	※3件の公報は親出願/子出願/孫出願の関係	セラジェクト、インコーポレイテッド
7	特開2013-027742		
8	特表2012-504160	多重薬物放出調節の可能なソリッドマイクロ構造体及びその製造方法	ヌリーエム ウェルネスカンパニーリミテッド
9	特開2019-005032	外皮内溶解性ニードル及びニードル装置	南郷 教裕

図 5 特許検索競技大会 2021 年の正解公報例

ChatGPT(GPT3.5)を利用した特許調査の具体例

ChatGPT(GPT3.5)への質問
下記内容の日本語の出願の類似特許を教えてください
「水溶性固形成分を基材とする経皮投与用マイクロニードルであって、前記マイクロニードルは内部に微粒子を含み、前記微粒子は乳酸ポリマー、グリコール酸ポリマー及び乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーと薬理活性物質を含有する経皮投与用マイクロニードル。」
ChatGPTの回答
以下は、上記の特許出願内容に類似する日本の特許出願例です：
特許出願番号JP2020-142164A - 「薬物貯蔵性マイクロニードル及びその製造方法、経皮送達装置」
特許出願番号JP2020-194018A - 「マイクロニードル、薬剤拡散システム、及びその製造方法」
特許出願番号JP2021-058903A - 「乳酸ポリマー、グリコール酸ポリマー及び乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーと、有効成分を含む経皮吸収用マイクロニードル及びその製造方法」
これらの特許出願は、経皮投与用マイクロニードルに関するものであり、乳酸及びグリコール酸ポリマーを含む微粒子を内部に含んでいます。また、薬理活性物質を含有し、皮膚透過性があるという点でも、上記の特許出願と類似しています。ただし、微粒子の具体的な構成や、その他の特徴は異なる場合がありますので、詳細に比較検討する必要があります。

商用CyberPatent Desk[®]による出願番号照会

No.	出願番号	公開・公表番号	発明の名称	出願人(最新)
1	特願2020-142164	特開2022-37822	熱交換器及び凝縮装置	合同会社電視工芸
2	特願2020-194018	特開2022-82870	遊技機	株式会社ニューギン
3	特願2021-58903	特開2022-155596	ロボット装置	株式会社アイシン

図 6 ChatGPT(GPT3.5) の回答 (2023 年 4 月時点)

特許リストは明らかに間違っておりまさに生成されていることが分かる。

2023 年 6 月時点の回答では間違った特許リストを生成することは無くなったが正解公報は 2021 年 9 月以前に公開されているのでデータ更新を理由にするのは不適切である。

ChatGPT(GPT4) の回答を検証すると 2023 年 6 月時点でも間違った特許リストを生成していた。

BingAI は最初に Bing で Web 検索を行い、次に下記回答を生成した。

ChatGPT(GPT3.5)を利用した特許調査の具体例

ChatGPT(GPT3.5)への質問 2023/06末時点
 下記内容の日本出願の類似特許を教えてください
 「水溶性固形成分を基材とする経皮投与用マイクロニードルで、あって、前記マイクロニードルは内部に微粒子を含み、前記微粒子は乳糖ポリマー、グリコール酸ポリマー及び乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーと薬理活性物質を含有する経皮投与用マイクロニードル。」

ChatGPTの回答
 申し訳ありませんが、私のデータベースは2021年9月までの情報しか持っていないので、最新の特許情報を提供することはできません。特許の申請や公開は時間がかかる場合があります。最新の特許情報を入手するには、特許庁や特許データベースなどの公式の情報源をご利用ください。

ChatGPTの回答内容

・正解公報9例は2021年9月以前に公開されているのでデータ更新を理由にするのは不適切

・特許庁や特許データベースなどを推奨することは適切

図7 ChatGPT(GPT3.5)の回答(2023年6月時点)

ChatGPT(GPT4)を利用した特許調査の具体例

ChatGPT(GPT4)への質問 2023/06末時点
 下記内容の日本出願の類似特許を教えてください
 「水溶性固形成分を基材とする経皮投与用マイクロニードルで、あって、前記マイクロニードルは内部に微粒子を含み、前記微粒子は乳糖ポリマー、グリコール酸ポリマー及び乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーと薬理活性物質を含有する経皮投与用マイクロニードル。」

ChatGPTの回答
 私の知識は2021年までのもので、それ以降の特許情報は提供できません。また、特定の特許の詳細やそれに関連する特許を調べるためには、具体的な特許番号や特許名が必要です。ただし、あなたが提供した情報に基づいて、類似の概念を持つ特許を概説します。

日本特許JP2006198036A:「マイクロニードルとその製造方法」この特許では、バイオデグラダブルマイクロニードルの製造方法について述べられています。このマイクロニードルは、薬物を保持し、皮膚組織に経皮投与するためのものです。

日本特許JP2014218169A:「マイクロニードルアレイ、マイクロニードルアレイの製造方法及びその使用方法」ここでは、生体適合性および生体内分解性を有するバイオデグラダブルマイクロニードルアレイについて記述されています。該マイクロニードルアレイは、人体に対して最小限の侵襲性で、治療または予防薬を効率よく送達するためのものです。

これらの特許は、あなたが述べた特許の内容と似たような概念を持つものですが、具体的な特許を探すには、特許データベースの検索が必要です。特許の検索と分析には専門的な知識が必要で、特許事務所や専門家の協力が一般的です。

図8 ChatGPT(GPT4)の回答(2023年6月時点)

ChatGPT(GPT4)を利用した特許調査の検証

商用CyberPatent Desk®による公開番号照会

出願番号 No.	筆頭IPC(最新) No.	発明の名称 No.	公報番号 No.	出願人(最新) No.
特開2005-10915	A63F7/02	遊技機	特開2006-198036 特許4444127	株式会社ニューギン
特開2013-98802	B60R25/34	姿勢検知システム、当該システムを構成する送信装置及び受信装置	特開2014-218166 特許6036528	株式会社デンソー

図9 ChatGPT(GPT4)の回答の検証

BingAIを利用した特許調査の具体例

BingAIへの質問 2023/06末時点
 下記内容の日本出願の類似特許を教えてください
 「水溶性固形成分を基材とする経皮投与用マイクロニードルで、あって、前記マイクロニードルは内部に微粒子を含み、前記微粒子は乳糖ポリマー、グリコール酸ポリマー及び乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーと薬理活性物質を含有する経皮投与用マイクロニードル。」

BingAIの回答
 水溶性固形成分を基材とする経皮投与用マイクロニードルで、あって、前記マイクロニードルは内部に微粒子を含み、前記微粒子は乳糖ポリマー、グリコール酸ポリマー及び乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーと薬理活性物質を含有する経皮投与用マイクロニードル」を検索しています

回答を生成しています…

水溶性固形成分を基材とする経皮投与用マイクロニードルで、あって、前記マイクロニードルは内部に微粒子を含み、前記微粒子は乳糖ポリマー、グリコール酸ポリマー及び乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーと薬理活性物質を含有する経皮投与用マイクロニードル」という特許出願に関する情報を提供しています。検索した結果、この特許出願に関連する情報は見つかりませんでした。もう少し具体的な情報があれば、お探しの情報に近いものを探せるかもしれません。何か他にお探しの情報があれば、お知らせください。

図10 BingAIの回答(2023年6月時点)

「検索した結果、この特許出願に関連する情報は見つかりませんでした。」

間違った特許リストを生成されるよりはよいが具体的な特許調査はこの件では行えなかった。

Google Bardを利用した特許調査の具体例

Google Bardへの質問 2023/06末時点
 下記内容の日本出願の類似特許を教えてください
 「水溶性固形成分を基材とする経皮投与用マイクロニードルで、あって、前記マイクロニードルは内部に微粒子を含み、前記微粒子は乳糖ポリマー、グリコール酸ポリマー及び乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーと薬理活性物質を含有する経皮投与用マイクロニードル。」

Google Bardの回答
 はい、日本語で答えます。

「水溶性固形成分を基材とする経皮投与用マイクロニードルで、あって、前記マイクロニードルは内部に微粒子を含み、前記微粒子は乳糖ポリマー、グリコール酸ポリマー及び乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーと薬理活性物質を含有する経皮投与用マイクロニードル。」

この特許は、水溶性固形成分を基材とする経皮投与用マイクロニードルに関するものです。マイクロニードルは、薬物を含有する微小な針で、皮膚に穿刺することで、薬物を直接体内に取り込むことができます。この特許では、マイクロニードルの内部に微粒子を含ませることで、薬物の吸収率を向上させています。微粒子は、乳糖ポリマー、グリコール酸ポリマー、乳酸グリコール酸コポリマーから選択される少なくとも1つのポリマーで構成されており、薬理活性物質を含有しています。

この特許に類似する特許としては、以下のようなものが挙げられます。

特許第5796346号「経皮投与用マイクロニードル」
 特許第5919656号「経皮投与用マイクロニードル」
 特許第8039013号「経皮投与用マイクロニードル」
 これらの特許は、すべて水溶性固形成分を基材とする経皮投与用マイクロニードルに関するものです。ただし、微粒子の構成や薬理活性物質が異なります。

図11 Google Bardの回答例1(2023年6月時点)

Google Bardは回答例1～回答例3の3種類の回答例を生成した。回答例1の特許リスト3件全て間違っていた。

回答例2は「申し訳ありませんが、その特許は見つけれませんでした。」

回答例3の特許リスト5件全て間違っていた。

残念ながら生成系AIのChatGPT、BingAI、Google Bardを利用しての直接的な特許調査は難しいことが良く分かった。ChatGPTに直接「ChatGPTは特許データベースの検索を行えますか?」と尋ねてみた。

ChatGPT 3.5の回答は、

「はい、ChatGPTは特許データベースの検索を行うことができますが、いくつかの制約があります。ChatGPTはテキストベースの質問応答モデルであり、インターネット上の情報や一般的な知識に基づいて回答を生成しますが、特許データベースへの直接アクセスは持っていません。」

一般的な特許情報や特許に関連するトピックについての質問には、一般的な知識や特許に関連する情報を提供できる可能性があります。特定の特許に関する情報を調べる必要がある場合は、特許データベースや専門の特許検索ツールを利用することをおすすめします。

特許検索に関連して具体的な質問があれば、お答えできる範囲でお手伝いいたしますので、どうぞお尋ねください。」

上記回答を踏まえてChatGPT、Bing AI、BardのWeb検索機能について検討した。



5 ChatGPT、Bing AI、Bard の Web 検索機能

Web 検索機能の検討対象として「スーパーコンピュータ「富岳」を使用した大規模言語モデルのプレスリリース」を使用した。このフレーズで Google 検索を行うと検索結果の上位から Fujitsu、理化学研究所、東京工業大学、cloud watch、pc watch、nikkei、ZDNet、東北大学のプレスリリースや記事がヒットする（2023年7月2日時点）。

ChatGPT、Bing AI、Google BardのWeb検索機能

入力プロンプト: スーパーコンピュータ「富岳」を使用した大規模言語モデルのプレスリリースについて教えてください。

GPT-3.5	<p>プレスリリースの例: FOR IMMEDIATE RELEASE 日付: 2023年7月2日 スーパーコンピュータ「富岳」を使用した大規模言語モデルの開発に成功 東京、日本 - 世界的なテクノロジー企業であるOpenAIは、最新のスーパーコンピュータ「富岳」を使用して開発された大規模な言語モデルに関する成果を発表しました。この革新的なプロジェクトは、人工知能分野における大きな飛躍となるものであり、自然言語処理や機械翻訳、文書生成など、さまざまな応用分野で大きなインパクトを持つことが期待されています。以下 省略</p>
GPT-4	<p>申し訳ありませんが、私の知識は2021年9月までの情報であり、それ以降の具体的な出来事については知識を持っておりません。したがって、2023年に富岳スーパーコンピュータで行われた大規模言語モデルに関するプレスリリースについての情報は提供できません。</p> <p>しかし、富岳のようなスーパーコンピュータを使って大規模言語モデルを訓練することは、大量の計算リソースを必要とする大規模なディープラーニングプロジェクトに非常に適しています。富岳は、多数のコアと高速な通信を備えたスーパーコンピュータであり、多くの並列処理タスクを同時に処理する能力があります。</p> <p>大規模な言語モデルは、非常に大きなデータセットから言語のパターンを学習するために、深層学習という形の人工知能を使用します。これは大量の計算力を必要とするため、スーパーコンピュータのような強力なマシンが必要となることが多いです。以下 省略</p>

図 12 ChatGPT3.5、ChatGPT4 の回答

ChatGPT3.5 の回答は一目でフェイクニュースとわかるレベルの生成された回答である。ChatGPT4 のデフォルトでは Web 検索機能を備えていないことが推定される回答である。ChatGPT で「GPT-4」を利用するためには有料版（月額 20 ドル）の「ChatGPT Plus」を利用する必要がある。無料版の ChatGPT では、言語モデルは「GPT-3.5」のみである。「ChatGPT Plus」は「GPT-3.5」「GPT-4」を選択できる。「GPT-4」は「デフォルト」、「Browse with bing (Model: Web Browsing)」、「Plugins」を選択できる。「Browse with bing」を選択すると bing による Web 検索機能が利用できる。またインターネット上の URL を指定してページにアクセスできる。「Plugins」を選択すると各種プラグインを選択できるようになる。「AskYourPDF」プラグインを組み込むと URL を指定して Web 上の PDF を読み込めるようになる。

図 13 に ChatGPT4(Web Browsing)、BingAI の回答を示す。

ChatGPT4(Web Browsing)、BingAI どちらも Web 検索を行い、検索結果から、Fujitsu のプレスリリースを選択して回答を生成している。この「富岳」を

ChatGPT、Bing AI、Google BardのWeb検索機能

入力プロンプト: スーパーコンピュータ「富岳」を使用した大規模言語モデルのプレスリリースについて教えてください。

GPT-4 Model: Web Browsing	<p>スーパーコンピュータ「富岳」を使用した大規模言語モデルの研究開発についてのプレスリリースは、2023年5月22日に東京工業大学、東北大学、富士通株式会社、および理化学研究所から発表されました。これらの組織は、「富岳」政策対応枠を利用して大規模言語モデル (Large Language Model, LLM) の分散並列学習手法の研究開発を2023年5月から実施しています。この研究開発の成果物を公開することで、アカデミアや企業が幅広く使える大規模言語モデルの構築環境を整え、国内におけるAIの研究力向上に貢献し、学術および産業の両面で「富岳」の活用価値を高めることを目指しています。以下省略</p>
BingAI	<p>京工業大学、東北大学、富士通、理化学研究所は、「富岳」政策対応枠において、スーパーコンピュータ「富岳」を活用した大規模言語モデル (Large Language Model, LLM) の分散並列学習手法の研究開発を2023年5月から実施しています。</p> <p>このプレスリリースについて詳しく知りたい場合は、以下のリンクを参照してください。</p> <p>「スーパーコンピュータ「富岳」政策対応枠における大規模言語モデルの分散並列学習手法の研究開発について」(https://pr.fujitsu.com/jp/news/2023/05/22.html)</p>

図 13 ChatGPT4(Web Browsing)、BingAI の回答

使用した「大規模言語モデル」の「プレスリリース」の例では、特徴的なキーワードの組み合わせで注目を集める内容であり正解を見つけるのは容易である。Web 検索結果を評価する場合、Web 検索の難易度と検索ユーザーが望んでいる内容との「適合」の観点から結果を評価する必要がある。「適合」の観点からは、ユーザーが望んでいるものを AI 側で判断するか、上位に表示することが望ましい。この検索ユーザーが望んでいる内容と「適合」しないものを AI が選択してしまうと生成する回答は不適切なものとなる。

なお、ChatGPT Plus のヘルプセンターのアナウンスでは「2023年7月3日の時点で、ChatGPT Browse ベータ版では、望ましくない方法でコンテンツが表示される場合があることがわかりました。細心の注意を払って Browse with Bing ベータ機能を無効にしました。私たちはベータ版をできるだけ早く戻すよう取り組んでいます。ご理解のほどよろしくお願いたします。」とのことである。

ChatGPT、Bing AI、Google BardのWeb検索機能

入力プロンプト: スーパーコンピュータ「富岳」を使用した大規模言語モデルのプレスリリースについて教えてください。

Google Bard	<p>2023年5月22日、理化学研究所は、スーパーコンピュータ「富岳」を使用した大規模言語モデル (LLM) のプレスリリースを発表しました。このLLMは、テキストとコードの膨大なデータセットでトレーニングされており、テキストを生成し、言語を翻訳し、質問に答えたり、コードを生成し、クリエイティブ コンテンツを作成したりすることができます。</p> <p>このLLMは、理化学研究所の計算科学研究センター (R-COS) と東京工業大学の研究チームが共同で開発しました。研究チームは、LLMをトレーニングするために、「富岳」の10ベタフロップス以上の計算能力を使用しました。</p> <p>このLLMは、さまざまなタスクで人間のレベルのパフォーマンスを達成できることが示されています。たとえば、テキストを生成するタスクでは、人間が生成したテキストと見分けがつかないテキストを生成することができます。また、言語を翻訳するタスクでは、人間が翻訳した翻訳と見分けがつかない翻訳を生成することができます。また、質問に答えるタスクでは、人間が答えた答えと見分けがつかない答えを出すことができます。</p> <p>このLLMは、さまざまな分野で使用されることが期待されています。たとえば、教育分野では、このLLMを使用して、生徒に個別に最適化された授業をすることができます。また、医療分野では、このLLMを使用して、患者の病状を診断したり、治療法を開発したりすることができます。また、ビジネス分野では、このLLMを使用して、顧客のニーズを理解したり、新しい製品やサービスを開発したりすることができます。</p> <p>このLLMは、LLMの将来の開発に大きな影響を与える可能性があります。LLMは、テキストの生成、言語の翻訳、質問への回答など、さまざまなタスクに使用することができます。このLLMにより、LLMはこれらのタスクでより正確で効率的に動作することができるようになりますと期待されます。</p>
-------------	---

図 14 Google Bard の回答

Google Bard の回答を図 14 に示す。一見すると正しそうであるがオリジナルのプレスリリースと比べると抜け漏れがあったり、逆にオリジナルにない事が回答されていたりでファクトチェックに手間がかかる回答である。そこで ChatGPT4 の「AskYourPDF」プラグインを利用して東京工業大学のホームページに掲載されているオリジナルのプレスリリース PDF を元の文章として、Google Bard の回答文章を対象文章として、対象文章の相違点を指摘するよう指示した。結果を図 15 に示す。

ChatGPTによる文章比較と相違点抽出

入力プロンプト: 下記URLの内容を元の文章として、次の対象文章と比較して、対象文章の相違点を指摘してください。
<https://www.titech.ac.jp/news/pdf/tokyotechpr20230522-fugakugpt.pdf> ←元の文章
 対象文章 (Google Bardの回答文章)

2023年5月22日、理化学研究所は、スーパーコンピュータ「富岳」を使用した大規模言語モデル(LLM)のプレスリリースを発表しました。以下省略

<p>対象文章と元のPDF内容との間には以下の相違点があります。</p> <ol style="list-style-type: none"> 1. 対象文章では、理化学研究所がプレスリリースを発表したと記述されていますが、元のPDFでは東京工業大学、東北大学、富士通、理化学研究所の4者が共同でプロジェクトを開始すると記述されています。 2. 対象文章では、LLMがテキストとコードの大量のデータセットで訓練され、テキスト生成、言語翻訳、質問応答、コード生成、クリエイティブコンテンツ作成などが可能であると記述されていますが、元のPDFではそのような詳細な機能については触れられていません。 3. 対象文章では、LLMが理化学研究所の計算科学研究センター(R-CCS)と東京工業大学の研究チームが共同で開発したと記述されていますが、元のPDFでは東京工業大学、東北大学、富士通、理化学研究所の4者が共同で開発を行うと記述されています。 4. 対象文章では、LLMが人間のレベルのパフォーマンスを達成できると記述されていますが、元のPDFではそのような記述はありません。 5. 対象文章では、LLMが教育、医療、ビジネスなどの分野で使用されることが期待されていると記述されていますが、元のPDFではそのような具体的な応用例については触れられていません。 6. 対象文章では、LLMがテキスト生成、言語翻訳、質問応答などのタスクでより正確で効率的に動作することが期待されていると記述されていますが、元のPDFではそのような記述はありません。

図 15 ChatGPT による文章比較と相違点抽出

回答結果は、特許調査の実務でも調査目的により十分に使えるレベルにあると予想されるものである。

EPO の Espacenet、商用 AI 特許調査ツールの amplified.ai¹¹⁾ のサイトで人がブラウザでアクセスすると表示される特定の特許を URL で指定して、ChatGPT4(Browse with Bing) にタイトルと要約の抽出を指示した回答を以下に示す。

「申し訳ありませんが、指定された URL からは情報を取得できませんでした。このウェブサイトは、私のブラウジングツールが情報を取得するのを阻止する robots.txt というファイルを使用しています。これはウェブサイトの所有者が特定のウェブクローラー(自動的にウェブページを閲覧するプログラム)からページの閲覧を制限するために使用する標準的な方法です。

他の情報源から必要な情報を見つけることが可能であれば、お知らせください。また、特許情報については、各国の特許庁のウェブサイトや Google Patents など、他の特許データベースを利用することも可能です。」と

「例えば、特許番号 JPWO2006016647A1 を Google Patents で検索することで、該当の特許についての情報を見つけることができます。

この方向で進めてよろしいでしょうか？」

「この方向で進めてください。」とお願いすると、

「ごめんなさい、一部のテクニカルな問題により特許情報を引用することができません。ですが、私が理解した内容をベースに説明を試みます。

この特許は、「微細針を用いた経皮的薬物投与装置」について説明しています。経口投与や注射投与は一般的な薬物投与方法ですが、これらには各々の問題点が存在します。経口投与では薬物の吸収が不完全であったり、食事との相互作用が問題となることがあります。また、注射投与は痛みを伴うことが多く、自己投与が難しいという問題があります。・・・中略

なお、上記の説明は特許の内容を理解した結果の説明であり、特許文書の正確な言葉を引用したものではありませんのでご了承ください。」とのことで直接的な引用はできないが「私が理解した内容」は概ね正しいと思われる。

6 ChatGPT のプロンプトエンジニアリング

プロンプトエンジニアリングとは、AI モデルに対して適切な質問や指示を与えることで、より望ましい結果を引き出す技術である。ChatGPT に質問したプロンプトエンジニアリングの基本についての回答を図 16 に示す。

ChatGPT4によるプロンプトエンジニアリング

ChatGPTのプロンプトエンジニアリングの基本について教えてください。

プロンプトエンジニアリングは、人間とAIモデルの間の対話を制御するための手法です。以下に、その基本的な手順と原則を示します:

1. ****モデルの理解**** AIモデルの機能と限界を理解することから始まります。たとえば、GPT系のモデルは大量のテキストデータから学習しており、一般的な知識と言語パターンについての広範な理解を持っていますが、個々のユーザーやリアルタイムの情報については知りません。
2. ****明確なプロンプト**** プロンプトは具体的で明確でなければならず、何を求めているのか、どのような形式で回答が欲しいのかを明確にする必要があります。たとえば、「天気は?」と尋ねるよりも、「ニューヨークの今日の一般的な天気はどうですか?」と具体的に尋ねる方がより良い結果を得られます。
3. ****コンテキストの使用**** GPT系のモデルは与えられたプロンプトの直前のコンテキストを利用します。したがって、必要な情報を含む十分なコンテキストを提供することが重要です。
4. ****試行とエラー**** プロンプトエンジニアリングはしばしば試行錯誤のプロセスです。異なるプロンプトを試し、結果を評価し、プロンプトを改善するというステップを繰り返します。
5. ****出力の制御**** プロンプトの形式はモデルの出力に影響を与えます。たとえば、質問形式のプロンプトは通常、簡潔な回答を生み出し、物語性のあるプロンプトはより詳細な回答を生み出す傾向があります。
6. ****指示的な言葉の使用**** あいまいなプロンプトよりも、具体的に指示的な言葉を使ったプロンプトの方が、望む結果を得やすいです。

これらの基本的な原則を理解し、それに基づいてプロンプトを作成することで、AIモデルから望む出力を得る確率を高めることができます。

図 16 ChatGPT によるプロンプトエンジニアリング



7 特許調査における大規模言語モデルの活用

ChatGPT4 に特許調査における大規模言語モデルの活用方法の可能性を5項目挙げてもらい、機械学習の専門家、ベテラン特許サーチャー、発明者の立場からのコメントを求め、最後に表にまとめた

特許調査における大規模言語モデルの活用方法

先ほどの特許調査における大規模言語モデルの活用方法の可能性5項目を挙げて簡単なコメントを付けてください。

ChatGPT4の回答
大規模言語モデルの活用に関して、発明者の視点からは以下の要点が重要になります。

1. ****特許文書の自動読解**** モデルの解釈が発明者の技術的な視点と一致しているか、または新しい視点を提示しているかを確認することが重要です。モデルが特許文書を正確に解釈し、発明者が理解しやすい形で情報を提示することが求められます。
2. ****技術トレンドの分析**** 発明者は新しいアイデアや製品の開発に関与しているため、その分野の最新のトレンドを把握することは非常に重要です。そのため、モデルが提供する技術トレンドの情報が、事実に基づいており、また最新の情報を反映していることが必要です。
3. ****類似特許の検索**** 自分の発明が既存の特許とどの程度類似しているかを理解することは、発明者にとって重要です。言語モデルが提供する類似特許のリストが正確であること、そして、自分の発明との関連性を明確に示すことが求められます。
4. ****侵害分析**** 自分の発明が他の特許を侵害していないことを確認することは、法的な問題を避けるために重要です。そのため、言語モデルが提供する侵害分析が正確であり、そして詳細な情報を提供することが求められます。
5. ****特許のドラフティング**** 発明者は自分のアイデアを適切に表現することが求められます。そのため、モデルが提供するドラフトは、発明の本質を捉え、そして適切な法的な形式で表現することが重要です。

これらの観点から、発明者は大規模言語モデルの活用を評価しますが、最終的には自分の発明の特性と目的を理解している専門家の助けが必要になるでしょう。

図 17 特許調査における大規模言語モデルの活用方法

特許調査における大規模言語モデルの活用方法

先ほどの特許調査における大規模言語モデルの活用方法の可能性5項目を挙げて簡単なコメントを付けてください。

ChatGPT4の回答
大規模言語モデルを特許調査に活用するためには、いくつかの重要な考慮事項があります。

1. ****特許文書の自動読解**** これを実現するためには、特許文書の構造を理解するためのアルゴリズムが必要です。特許文書は特定の形式と用語を使用するため、モデルはその特性を学習する必要があります。それは教師あり学習を用いた分類問題や、特許文書の自動要約の生成などのタスクを通じて可能となります。
2. ****技術トレンドの分析**** これは、時間に基づく情報の分析能力が必要となります。トピックモデリングや時系列分析を用いて、特許データセット内の技術トレンドを抽出することが可能です。
3. ****類似特許の検索**** 文章の意味的類似性を計算するために、自然言語処理(NLP)の手法、特にベクトル空間モデルやドキュメントモデリングを使用します。類似度の評価は、コサイン類似度やjaccard類似度等のマトリクスを用いて行うことができます。
4. ****侵害分析**** 侵害分析は比較的複雑であり、法的な解釈も必要となります。そのため、特許クレームの解析には、モデルが法的な文脈や、特許クレームの形式と表現に精通していることが必要となります。
5. ****特許のドラフティング**** 言語モデルは、事前に学習した文章のスタイルや形式に従ってテキストを生成する能力を持つため、特許のドラフティングには有用です。ただし、重要なものは、生成されたテキストが法的な要件を満たしているかどうかをチェックするための人間のレビューが必要であるという点です。

これらのタスクはすべて、大規模な特許文書データセットを用いたトレーニングが必要となります。また、各タスクは異なるタイプの教師あり学習や教師なし学習を必要とするため、その実装は複雑な問題となります。最後に、各タスクの結果の評価には、人間のエキスパートによるレビューが必要となるでしょう。

図 18 機械学習の専門家の立場からの回答

特許調査における大規模言語モデルの活用方法

先ほどの特許調査における大規模言語モデルの活用方法の可能性5項目を実現させるためにベテラン特許サーチャーの立場からコメントしてください。

ChatGPT4の回答
大規模言語モデルを特許調査に活用するためには、以下の考慮事項が特許サーチャーの視点から重要となります。

1. ****特許文書の自動読解**** モデルが適切に特許文書を解釈しているかを確認するためには、モデルの出力とサーチャー自身の解釈を定期的に比較することが必要です。また、モデルの解釈が法的に有効であり、技術的な観点からも正確であることを確認することが重要です。
2. ****技術トレンドの分析**** トrend分析の結果を適切に解釈するためには、サーチャーがその分野の専門知識を持つことが重要です。モデルはあくまでトレンドを示すのみで、それが何を意味するのか、またそれが事業や特許戦略にどのように影響するのかを理解するためには、深い専門知識が必要です。
3. ****類似特許の検索**** モデルが提供する類似特許のリストを検証する際、サーチャーはそのリストが完全であること、また正確であることを確認する必要があります。それには、特許分野に対する深い理解と、サーチャー自身の専門知識が求められます。
4. ****侵害分析**** モデルが侵害分析を行う際には、法的な知識だけでなく、実際の製品やプロセスの理解も必要となります。このため、サーチャーはモデルの出力を、具体的な製品やプロセスの詳細な知識に基づいて評価する必要があります。
5. ****特許のドラフティング**** 特許文書のドラフトは、法的な要件だけでなく、技術的な正確性も必要とするため、モデルによって生成されたドラフトをサーチャーが検証することが不可欠です。

総じて、大規模言語モデルは特許調査のプロセスを助けるツールとして有用ですが、それらのモデルが出力する結果は、ベテランの特許サーチャーによる定期的な検証と評価が必要です。

図 19 ベテラン特許サーチャーの立場からの回答

特許調査における大規模言語モデルの活用方法

先ほどの特許調査における大規模言語モデルの活用方法の可能性5項目を実現させるために発明者の立場からコメントしてください。

ChatGPT4の回答
大規模言語モデルの活用に関して、発明者の視点からは以下の要点が重要になります。

1. ****特許文書の自動読解**** モデルの解釈が発明者の技術的な視点と一致しているか、または新しい視点を提示しているかを確認することが重要です。モデルが特許文書を正確に解釈し、発明者が理解しやすい形で情報を提示することが求められます。
2. ****技術トレンドの分析**** 発明者は新しいアイデアや製品の開発に関与しているため、その分野の最新のトレンドを把握することは非常に重要です。そのため、モデルが提供する技術トレンドの情報が、事実に基づいており、また最新の情報を反映していることが必要です。
3. ****類似特許の検索**** 自分の発明が既存の特許とどの程度類似しているかを理解することは、発明者にとって重要です。言語モデルが提供する類似特許のリストが正確であること、そして、自分の発明との関連性を明確に示すことが求められます。
4. ****侵害分析**** 自分の発明が他の特許を侵害していないことを確認することは、法的な問題を避けるために重要です。そのため、言語モデルが提供する侵害分析が正確であり、そして詳細な情報を提供することが求められます。
5. ****特許のドラフティング**** 発明者は自分のアイデアを適切に表現することが求められます。そのため、モデルが提供するドラフトは、発明の本質を捉え、そして適切な法的な形式で表現することが重要です。

これらの観点から、発明者は大規模言語モデルの活用を評価しますが、最終的には自分の発明の特性と目的を理解している専門家の助けが必要になるでしょう。

図 20 発明者の立場からの回答

特許調査における大規模言語モデルの活用方法

先ほどの特許調査における大規模言語モデルの活用方法の可能性5項目を機械学習の専門家、ベテラン特許サーチャー、発明者の立場からのコメントの重要ポイントを表にまとめてください。

ChatGPT4の回答
以下の表は、大規模言語モデルの特許調査における活用可能性について、機械学習の専門家、ベテラン特許サーチャー、そして発明者の立場からの重要な視点をまとめたものです。

活用可能性の項目	機械学習の専門家	ベテラン特許サーチャー	発明者
特許文書の自動読解	アルゴリズムによる特許文書の構造の理解が必要	モデルの出力とサーチャー自身の解釈を定期的に比較することが重要	モデルの解釈が発明者の技術的な視点と一致しているか確認する
技術トレンドの分析	時間に基づく情報の分析能力が必要	ベクトル空間モデルやドキュメントモデリングの活用が必要	最新の情報を反映した正確な技術トレンドの情報が重要
類似特許の検索	法的な解釈と特許クレームの形式と表現の理解が必要	類似特許のリストが正確であることを検証する必要がある	自分の発明との関連性を明確に示す正確な類似特許の情報が必要
侵害分析	モデルが侵害分析を行う際には、法的な知識だけでなく、実際の製品やプロセスの理解も必要となります	モデルの出力を、具体的な製品やプロセスの詳細な知識に基づいて評価する必要があります	自分の発明が他の特許を侵害していないことを確認する必要がある
特許のドラフティング	モデルによって生成されたドラフトをサーチャーが検証することが不可欠	モデルが提供するドラフトが発明の本質を捉えていることが重要	モデルが提供するドラフトが発明の本質を捉えていることが重要

以上の観点はそれぞれの立場から見た重要なポイントを示していますが、全ての利害関係者が連携して、モデルの改善と活用方法の最適化を進めていくことが重要です。

図 21 各立場のまとめ

8 AI 利用特許調査ツールを用いた特許調査

AI 利用特許調査ツールを用いた特許調査の例として特許文書の類似文書検索を Patentfield⁹⁾、パナソニックソリューションテクノロジー株式会社が提供している PatentSQUARE¹⁰⁾ を利用して検討した。

Patentfield のセマンティック検索の正式名称は AI セマンティック類似検索であり AI 検索に分類される。同じセマンティック検索という名称でも商用データベースによっては概念検索の場合もあり注意を要する。

Patentfield のセマンティック検索は内部では fastText (Facebook AI Research が開発した自然言語処理を高速、高精度化するのに有効なライブラリ) の分散表現ベクトルを使用しており収録されている全分野にわたって類似語を学習している。Patentfield の全文検索は明細書の検索対象を選択してキーワードにより検索できる。セマンティックスコアを使用することで、探したい任意の自然文書または特定の文献から類似度が

高い順にスコアリングされ、教師データを用意せずに探したい技術内容に近い特許文献から調査することができる。セマンティックスコアはブーリアン演算によるコマンド検索の母集団に組み合わせることも可能で、任意のキーワード、文書、特許番号などを指定することにより検索結果の母集団を変更することなく、容易に近い内容の技術から効率的に調査を行うことができる。

Patentfieldのブーリアン検索+AIセマンティック検索

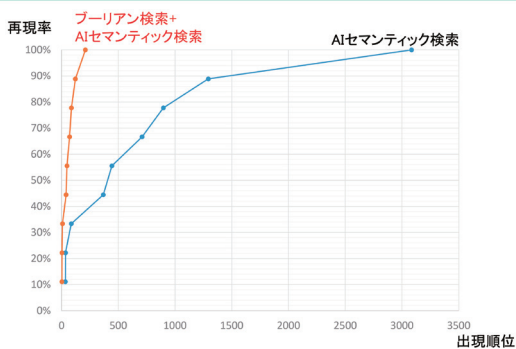


図 22 Patentfield のブーリアン×AI セマンティック検索

Patentfield のブーリアン検索×AI セマンティック検索画面を図 22 に示す。

ブーリアン検索で FI: マイクロニードル +F ターム : 微粒子×ポリエステル の検索を行った。AI セマンティック検索には請求項 1 を使用した。

Patentfieldのブーリアン検索+AIセマンティック検索



ブーリアン検索 (FI: マイクロニードル +F ターム : 微粒子×ポリエステル)
セマンティック検索範囲
SE: 名称/要約/請求の範囲/明細書/審査官フリーワード

図 23 Patentfield のブーリアン検索×AI セマンティック検索

AI セマンティック検索単独と「ブーリアン検索×AI セマンティック検索」結果を図 23 に示す。「ブーリアン検索×AI セマンティック検索」と組み合わせることで大幅に出現順位が向上している。

PatentSQUAREのAI検索事例



図 24 PatentSQUARE の AI 検索事例

PatentSQUARE の AI 検索画面を図 24 に示す。

PatentSQUAREのAI検索事例

PatentSQUAREのAI検索結果

No	AIスコア	出願番号	出願日	公開・公表公開・公表登録番号	登録日	出願人・特許権者の名称	特許番号
1	83.4	P2017-504	2015.04.22	P2017-524	2017.09.28, 2016.01.28	エルジー レチノール	1
2	83	P2011-534	2009.10.01	P2012-504	2012.02.16, 2010.04.08	ヌリーエム多量薬物	1
30	81.4	P2015-224	2015.11.11	P2016-084	2016.05.22, P6169149	2017.07.01 スモール マイクロニ	1
332	77.5	P2008-534	2006.09.06	P2009-504	2009.02.24, P5161776	2012.12.21 セラジェク薬物粒子	1
387	77.3	P2011-284	2011.12.24	P2012-114	2012.06.14	セラジェク薬物粒子	1
4086	72.6	P2012-234	2012.10.22	P2013-024	2013.02.07	セラジェク薬物粒子	1
7057	71	P2017-554	2016.04.06	P2018-514	2018.04.15, P6875289	2021.04.24 エルジー 難溶性薬	1

PatentSQUAREのAI再検索結果

No	AIスコア	出願番号	出願日	公開・公表公開・公表登録番号	登録日	出願人・特許権者の名称	特許番号
1	100.0	P2017-504	2015.04.22	P2017-524	2017.09.28, 2016.01.28	エルジー レチノール	1
2	99.2	P2011-534	2009.10.01	P2012-504	2012.02.16, 2010.04.08	ヌリーエム多量薬物	1
5	96.8	P2008-534	2006.09.06	P2009-504	2009.02.24, P5161776	2012.12.21 セラジェク薬物粒子	1
28	78.2	P2012-234	2012.10.22	P2013-024	2013.02.07	セラジェク薬物粒子	1
124	18.9	P2017-554	2016.04.06	P2018-514	2018.04.15, P6875289	2021.04.24 エルジー 難溶性薬	1
331	14	P2011-284	2011.12.24	P2012-114	2012.06.14	セラジェク薬物粒子	1
716	12	P2015-224	2015.11.11	P2016-084	2016.05.22, P6169149	2017.07.01 スモール マイクロニ	1

図 25 PatentSQUARE の AI 検索結果

PatentSQUARE の AI 検索結果を図 25 に示す。上段は AI 検索結果で正解情報の No1、No2 を選択して再検索したのが下段である。上段と下段の No (順位) を比べると下段の再検索結果の正解情報の順位が上がっているのが分かる。7 番目の正解情報の順位が 7057 位から 716 位になっておりこの検索集合を全てユーザーが適合判定すると仮定すると、読むべき情報数は 7057 件から 716 件に減少し、効率はおおよそ 10 倍になる。

これはいわゆる適合性フィードバックに基づくユーザーの検索意図を反映させた特許情報を再検索していることに相当する。ポイントは、ユーザーの検索意図を〇×をつけることで適合文書と不適合文書を明示的に検索システムにフィードバックしているところである。

PatentSQUARE の AI 検索結果 (ヒートマップ) を図 26 に示す。「検索テキスト位置表示ボタン」を押すと検索クエリ文書の位置を示す星印が表示される。表示

されているドットはクラスターになっておりドットをマウスでクリックするとドットのクラスターに属する公報がリスト表示される。

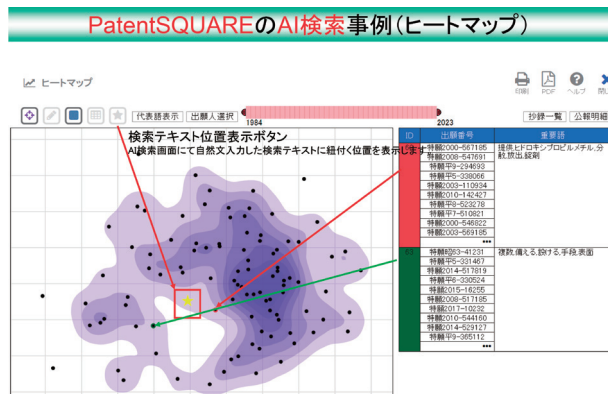


図 26 PatentSQUARE の AI 検索結果 (ヒートマップ)

9 BERT を用いた特許調査検討

BERT (Bidirectional Encoder Representations from Transformers) は、2018 年 10 月に Google の Jacob Devlin らの論文で発表された自然言語処理モデルである。

2019 年 12 月 2 日に株式会社日立ソリューションズより「活文 知的情報マイニング」の AI エンジンで強化し、高精度のテキスト分類を実現¹²⁾したシステムが販売開始されている。

「トランスフォーマーを用いた特許審査支援の探究」¹³⁾、「特許審査官が垣間見た深層学習 (Deep Learning) による自然言語処理の変革期」¹⁴⁾に BERT による特許調査に関して紹介されている。

BERT および BERT からの派生技術の「特許庁における AI 技術の活用の現状と最新の取組」¹⁵⁾は大変興味深い。

2020 年 11 月 21 日に米国およびその他の国の 1 億を超える特許出版物について、オープンソースツールを使用して BERT モデルをトレーニングする手法を概説したホワイト ペーパー¹⁶⁾がリリースされた¹⁷⁾。このホワイト ペーパーでは、特許出願の新規性を判断するための先行技術調査をより効果的に行う方法、特許分類を補助する分類コードを自動的に生成する方法、オートコンプリートなど、多くのユースケースでトレーニング済みモデルを使用する方法について説明している。このホワイトペーパーには、GitHub でホストされている

トレーニング済みモデルとともに、Colab ノートブックが付属している。

Japio 知財 AI 研究センターから SDGs 技術の見える化¹⁸⁾、特許情報を活用した SDGs の「見える化」¹⁹⁾が紹介されており BERT が分類に使われている。

文献^{20,21)}を参考にして BERT による特許調査の効率化検討を行った。より具体的には BERT を用いた特許文書の分類、類似文書検索、俯瞰可視化を検討した。

参考文献 20,21 のライブドアコーパスで BERT の取り扱いを検討して特許公報データを扱えるようにプログラムを追加・修正した。BERT による文書ベクトルの可視化を図 27 にしめす。

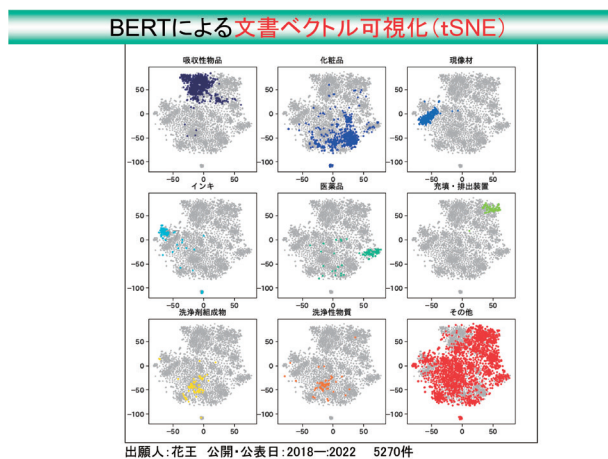


図 27 BERT による文書ベクトル可視化

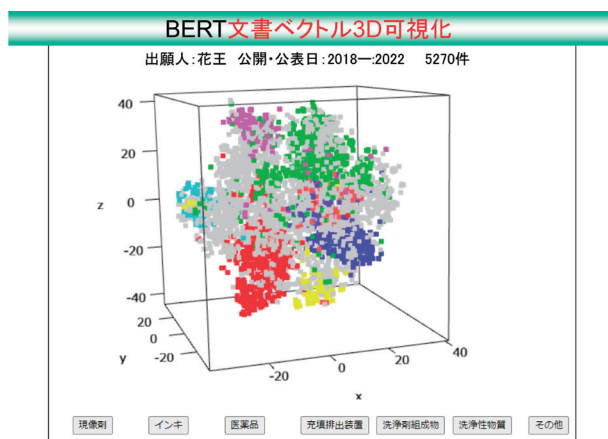


図 28 BERT による文書ベクトル 3D 可視化

図 28 はブラウザで表示され、マウス操作で 3D 表示の回転、拡大縮小、各カテゴリーのドットの ON / OFF 表示が可能である。

10 OSS を用いた特許調査

オープンソースを用いた機械学習の特許調査への応用に必要な構成要素を図 29 にまとめた。ChatGPT の API や機能を拡張するライブラリに関するに関する情報²²⁾も増えており OSS のライブラリを用いた各種プログラミングも ChatGPT と対話しながら行うことで格段に敷居が低くなっている。例えば、LlamaIndex は大規模言語モデルと外部データ（自分のデータ）を接続するためのライブラリである。LangChain は、ChatGPT などの大規模言語モデルの機能を拡張できるライブラリである。

オープンソースを用いた機械学習の特許調査への応用

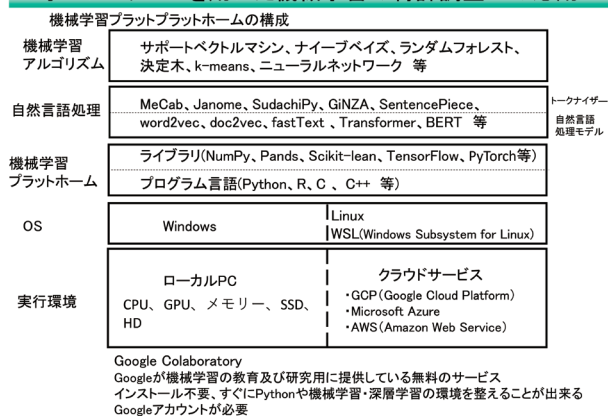


図 29 オープンソースを用いた機械学習の応用

7月上旬に ChatGPT の「Code Interpreter」が使用できるようになった。「Code Interpreter」とは、ChatGPT Plus のベータ版の機能で、ChatGPT 上で Python を使ったコードの実行や、ファイルのアップロード・ダウンロードができるようになる。例えば特許データベースからダウンロードした csv や Excel ファイルをアップロードして、自然言語による指示や Python コードを使用した各種作業結果を csv などでダウンロードすることが可能である。

ChatGPT に訊いた、ChatGPT の「Code Interpreter」でできること、できないことを図 30 に示す。

ChatGPTの「Code Interpreter」でできること、できないこと

ChatGPTの「Code Interpreter」についてできること、できないことを表形式にして日本語でExcelファイルに記入して、教えてください。

できること	できないこと
算術演算	インターネットへのアクセス(API呼び出し、Webスクレイピング等)
変数の割り当てと使用	外部ライブラリのインストール
データ構造(リスト、辞書など)の使用	システムコマンドの実行
関数の定義と呼び出し	データベースへの接続
クラスの定義と使用	GUIアプリケーションの作成
モジュールのインポート(標準ライブラリのみ)	マルチスレッド、マルチプロセスの実行
文字列、数値、日付等の操作	サブプロセスの作成
ファイルの読み書き('/mnt/data' ディレクトリ内のみ)	長時間実行するタスク(タイムアウト制限がある)
データの視覚化(matplotlib, seaborn等を使用)	
データ分析と機械学習のためのライブラリ(pandas, numpy, scikit-learn等)の使用	
LaTeXを使用した数式の表示	

図 30 ChatGPT の「Code Interpreter」

11 まとめ

大規模言語モデルによる特許調査の効率化検討を最近非常に注目されている ChatGPT に代表される生成系 AI に使われている「GPT」、既に自然言語処理分野で実績のある「BERT」を中心に紹介した。利用するにあたっての「BERT」、「GPT3」、「GPT4」の違いを図 31 にまとめた。現時点では ChatGPT、Bing AI、Bard をクエリを入力して直ぐに結果を求める直接的な特許調査に利用するには問題が多い。ただプロンプト次第で今まで非常に難しかったことが簡単に行える可能性が大きいのも確かである。

「BERT」、「GPT-3」、「GPT-4」の違い

	BERT	GPT-3.5	GPT-4
得意分野	意図推定、文章分類、質問応答、感情分析、文章要約、機械翻訳	要約、対話生成等の文章の生成	文章の生成(画像入力対応)
ユーザー側のチューニング方法	ファインチューニング タスクに応じた数百レベルの学習データ	Zero-shot、One-shot、Few-shot学習を活用して、タスクに適應	
利用形態	ソースプログラムを公開 事前学習済モデル公開 ユーザー側でファインチューニング	プログラム非公開 クラウドサービスのAPI GPT(有料版API、無料枠あり)	

BERT、GPT-3、GPT-4は、それぞれ異なる特徴と利用方法を持つ自然言語処理(NLP)モデル
BERT (Bidirectional Encoder Representations from Transformers)
GPT(Generative Pre-trained Transformer)

図 31 「BERT」、「GPT3」、「GPT4」の違い

現在は複数の AI 利用ツールが商用サービスとして提供されており選択肢は増えている。多くの場合、トリアルもでき、実際の実務データを使用してのツールの性能評価も可能である。調査目的に適合したツールを選択、あるいは構築することがポイントである。人間知能 HI (Human Intelligence) と人工知能 AI (Artificial Intelligence) の特徴把握と役割分担が重要である。



12 終わりに

本報告は 2023 年度の「アジア特許情報研究会」のワーキングの一環として報告するものである。

研究会のメンバーの皆様には様々な協力をしていただきました。ここに改めて感謝申し上げます。

参考文献

- 1) 岡崎直観. 大規模言語モデルの驚異と脅威
https://speakerdeck.com/chokkan/20230327_riken_llm
- 2) 岡野原大輔. 大規模言語モデルは新たな知能か ChatGPT が変えた世界
<https://www.iwanami.co.jp/book/b625941.html>
- 3) 太田邦史. 生成系 AI (ChatGPT, BingAI, Bard, Midjourney, Stable Diffusion 等) について
<https://utelecon.adm.u-tokyo.ac.jp/docs/20230403-generative-ai>
- 4) OpenAI の ChatGPT
<https://chat.openai.com/>
- 5) Microsoft BingAI
<https://www.bing.com/>
- 6) Google Bard
<https://bard.google.com/>
- 7) 特許検索競技大会 2021 年過去問
https://japio.or.jp/service/service04_05.html
- 8) CyberPatent Desk
<https://s.patent.ne.jp/>
- 9) Patentfield
<https://patentfield.com/>
- 10) PatentSQUARE
<https://www.panasonic.com/jp/business/its/patentsquare.html>
- 11) amplified.ai
<https://www.amplified.ai/ja/>
- 12) 「活文 知的情報マイニング」の AI エンジンで「BERT」で強化し、高精度のテキスト分類を実現
<https://www.hitachi-solutions.co.jp/company/press/news/2019/1125.html>
- 13) 前原義明 tokugikon 2020.5.29. no.297
トランスフォーマーを用いた特許審査支援の探究
<http://www.tokugikon.jp/gikonshi/297/297kiko5.pdf>
- 14) 石川雄太郎 tokugikon 2020.9.25. no.298
特許審査官が垣間見た深層学習 (Deep Learning) による自然言語処理の変革期
～ Japio 特許情報研究所での業務を通じて～
<http://www.tokugikon.jp/gikonshi/298/298tokusyu2.pdf>
- 15) 伊藤孝佑ら. 特許庁における AI 技術の活用 の現状と最新の取組
情報の科学と技術 . 2023, vol.73, no.7, p.256-261
https://doi.org/10.18919/jkg.73.7_256
- 16) Rob Srebrovic, Jay Yonamine
Leveraging the BERT algorithm for Patents with TensorFlow and BigQuery
https://services.google.com/fh/files/blogs/bert_for_patents_white_paper.pdf
- 17) AI、具体的には BERT が特許業界にどのように役立つか
<https://cloud.google.com/blog/ja/products/ai-machine-learning/how-ai-improves-patent-analysis/>
- 18) Japio 知財 AI 研究センター
SDGs 技術の見える化
<https://transtool.japio.or.jp/work/sdg/>
- 19) Japio 知財 AI 研究センター
特許情報を活用した SDGs の「見える化」
<https://transtool.japio.or.jp/work/data/sdg.pdf>
- 20) 近江崇宏ら オーム社 2021 年 200p
BERT による自然言語処理入門 :Transformers を使った実践プログラミング
- 21) 石田基広 シーアンアール研究所 2022 年 303p
Python で学ぶ テキストマイニング入門
- 22) 布留川英一 ボーンデジタル 2023 年 304p
OpenAI GPT-4/ChatGPT/LangChain 人工知能プログラミング実践入門

