

ChatGPT 4o

共有する

Hi

**あなた:**

特許を対象とするスコアリングモデルおよび大規模言語モデルの研究.docx  
ドキュメント

nonaka.pdf

PDF

講演の書き起こしファイルと講演に使用した資料です。要約を作成してください。

**ChatGPT:**

ChatGPT

**要約: 特許を対象とするスコアリングモデルおよび大規模言語モデルの研究 (野中尋史准教授)**

**発表の概要**

野中准教授は、特許を対象とするスコアリングモデルと大規模言語モデルの組み合わせにより、技術創造を促進する AI の開発について説明しました。重要技術の特定、スコアリング手法、特許権利期間の予測、技術の融合可能性の評価を中心に発表が進められました。

**主なポイント**

**1. 特許のスコアリング:**

- 特許の価値を「権利期間」や「引用情報」から定量的に評価する研究が行われています。
- 長期的な特許権利期間を予測するために、ディープラーニングを用いたニューラルハザードモデル (DeepSurv) を導入し、権利期間の延長可能性を見極めます。

**2. 引用ネットワークの成長性スコア:**

- 従来の引用数のみに依存する手法では過去の特許を過大評価するため、成長性を考慮する手法 (PageRank の改良版) を導入。
- 成長性スコアを IPC サブクラスごとに算出し、現在の技術発展において重要な特許を特定するモデルを開発しました。

**3. 技術の融合可能性の評価:**

- 異分野技術の融合により革新が生まれる可能性を重視し、データセットの類似性と課題の共起ネットワークを活用。
- 情報系データを基に、技術同士が融合可能かどうかを評価する手法を研究。

**4. 大規模言語モデルの活用:**

- 技術創造を目的に、知識グラフを利用して大規模言語モデルが因果関係に基

づく推論を行えるようデータセットを整備。

- 特許のテキストデータのみならず、画像データも含めて、マルチモーダル解析が可能なモデル構築を進行中。

#### 5. データセットの整備の重要性:

- 大規模言語モデルの発展には、良質なデータセットの整備が不可欠であり、特に官民学連携によるデータセット整備の重要性が強調されました。

この研究は、特許実務のみならず研究開発やマーケティング、戦略策定など幅広い分野に貢献する可能性を秘めています。

4o

あなた:

大規模言語モデルの活用:についてくわしく

**ChatGPT:**

ChatGPT

#### 大規模言語モデルの活用に関する詳細

野中准教授の研究において、大規模言語モデル (LLM) を活用した技術創造の実現には、以下のような戦略と方法論が採用されています。

#### 1. 知識グラフによる推論の強化

- 特許に含まれる情報 (技術的な課題、解決手段、効果) を単なるテキストデータとして扱うのではなく、「知識グラフ」に構造化して整理することが研究の中心です。
- 知識グラフを利用すると、課題と解決手段の因果関係が明示されるため、大規模言語モデルが推論を行う際の精度が向上します。例えば、「どの技術がどの課題を解決するのか」といった因果関係が明確にモデル内で扱えるようになります。

#### 2. 技術融合に向けた AI 支援の可能性

- 大規模言語モデルは、異なる技術分野における課題と解決策の相互関係を探り、技術の組み合わせによって生まれるイノベーションを提案できる可能性を持ちます。
- 例えば、機械工学と化学の技術を組み合わせることで、新しい発明や改善が見込まれる場合、このモデルが融合を提案し、応用範囲の広がりを示唆します。

#### 3. 画像データとテキストデータの統合 (マルチモーダル解析)

- 特許にはテキスト情報だけでなく、図面や画像が含まれることが多く、これらも技術の理解や解析において重要です。
- 野中准教授の研究グループでは、画像データとテキストデータの関連性を示すデータセットを整備し、画像と言語を同時に解析可能なマルチモーダルモデルを構築することを目指しています。これにより、画像とテキストを組み合わせた高度な技術創造や検索が実現できます。

#### 4. データセットの整備とその重要性

- 大規模言語モデルが有効に機能するためには、質の高いデータセットが不可欠です。

特に特許のような分野では、適切なデータセットが整備されていないことが課題です。

- 野中准教授の研究では、特許の図面とキャプションを関連付ける、課題と解決手段の関係を示す、といった知識グラフの整備を進めています。このようなデータセットは、モデルが技術情報を深く理解し、正確な推論を行うための基盤となります。

## 5. 推論精度の向上と誤情報の低減

- 一般的な LLM は自然言語の曖昧性に起因して誤情報を生成する可能性があります。知識グラフのような構造化データを取り入れることで、推論時の曖昧性を抑制し、精度の高い結果が得られるよう工夫されています。
- これにより、技術創造における推論結果の信頼性が向上し、実務や意思決定における利用価値が高まります。

## 6. 実務分野での応用

- この研究に基づいた LLM は、研究開発部門や知財部門、さらには経営層にとっても価値のあるツールとして期待されています。特に、特定の課題に対してどの技術を用いれば解決可能かを迅速に提案するなど、実務に直接応用できる可能性が高いです。
- また、マーケティングや商品開発の分野でも、LLM が市場ニーズに合致した技術提案を支援することが期待されます。

### まとめ

大規模言語モデルを特許情報処理に応用することで、特許分野や異分野の技術情報を体系的に解析し、新たな技術やアイデアを自動生成・推論する道が開かれます。この研究におけるモデルは、知識グラフの整備や画像データの統合により、より精緻な技術創造支援を目指しており、官民学の連携を通じてさらに発展することが期待されています(特許を対象とするスコアリングモデルおよび大規模…)(nonaka)。

40

**あなた:**

データセットの整備の重要性:をくわしく

**ChatGPT:**

ChatGPT

### データセットの整備の重要性に関する詳細

野中准教授の研究において、大規模言語モデル (LLM) を用いた特許情報処理を効果的に行うためには、質の高いデータセットの整備が極めて重要であると強調されています。この重要性について、以下の観点から説明します。

#### 1. LLM の性能向上に不可欠なデータセット

- 大規模言語モデルは、膨大で多様なデータセットによって学習されるため、モデルが推論や技術創造を行う際の精度や信頼性はデータの質に強く依存します。

- 特許情報は技術的な課題と解決手段、技術要素間の因果関係などが含まれ、知識が複雑かつ多層的です。これらの情報を正確に反映するデータセットがなければ、LLM は特許技術に関する適切な推論を行うのが困難になります。

## 2. 知識グラフの構築

- 特許情報を「知識グラフ」として構造化することで、特許技術に関する因果関係（課題→解決手段→効果）を明示的に表現できます。
- 知識グラフは、特許の課題、手段、技術的な要素、応用分野といったデータをネットワーク状に整理し、LLM が精度の高い推論を行えるよう補助します。知識グラフがあれば、技術的背景を踏まえた分析や、課題解決のための技術提案が可能になります。

## 3. マルチモーダルデータセットの整備

- 特許にはテキストだけでなく、図面や画像が含まれます。これらの視覚データは技術内容の理解を深めるために重要です。
- 野中准教授の研究では、図面と説明文の関係を明確にしたマルチモーダルデータセットを構築し、画像と言語情報を統合して解析するモデルの開発が進められています。これにより、LLM がテキストと画像の両方を考慮して技術内容を理解し、より高度な技術推論が可能になります。

## 4. 技術分野別のデータ整備と適用の差異

- 特許情報には、電気、機械、医薬品など異なる技術分野が含まれ、それぞれに異なる特徴があります。例えば、電気分野では引用数が多くなりやすい傾向がある一方で、医薬品分野では引用が少なくても影響力が高いことがあります。
- 各分野の特性を反映したデータセットを整備することで、分野特有の特徴を考慮した推論が可能になり、各分野に応じた適切な技術評価や提案が行えます。

## 5. オープンデータと官民学連携の必要性

- 野中准教授は、特許データセットの整備には官民学連携が必要であると述べています。現在、特許情報は一部オープンデータとして提供されていますが、最新のデータセットが不足しているのが現状です。
- 官民学が協力して特許データセットを充実させることで、LLM を用いた技術解析や技術創造の可能性が広がり、特許実務や研究においてもメリットが期待されます。例えば、特許情報がデータセットとして整備されると、研究者や企業が特許情報を迅速かつ効率的に利用できるようになります。

## 6. 2020 年代のデータセット整備の停滞と課題

- 野中准教授は、特許データセットの整備が 2010 年代以降停滞していると指摘しています。データセットの整備には大規模な人材やコストが必要で、持続的に整備される仕組みが確立されていないことが原因とされています。
- この課題を解決するためには、研究者だけでなく特許実務者を含めた広範な協力体

制が必要です。実務者が特許実務に役立つデータセットを整備することで、特許業務や研究活動の両方にとって価値のあるデータ基盤が形成されます。

## 7. データセット整備の今後の展望

- 野中准教授の研究では、官民学連携による特許データセットの充実が今後の課題解決につながると期待されています。特に、産業界と学术界が協力し合うことで、特許分野におけるデータセットがより高度化し、LLM による推論精度が高まることが予測されます。
- 最終的には、日本の技術や特許データの解析が世界的に進化し、特許解析や技術創造支援の分野で国際的に競争力を持つ可能性があります。

### まとめ

大規模言語モデルの有効な活用には、特許分野に特化したデータセットの整備が不可欠です。知識グラフ、マルチモーダルデータ、分野特性を考慮したデータなど、整備されたデータセットは、LLM の精度を飛躍的に高め、特許情報の理解や活用における実務上の支援だけでなく、技術創造の可能性も広げる基盤となります(特許を対象とするスコアリングモデルおよび大規模…)(nonaka)。

40

### あなた:

この研究は、特許実務のみならず研究開発やマーケティング、戦略策定など幅広い分野に貢献する可能性を秘めています。をくわしく

### ChatGPT:

ChatGPT

### 特許データと大規模言語モデル (LLM) を活用した研究の多分野への貢献可能性

野中准教授の研究で開発されている特許データ処理および技術創造 AI は、特許実務の効率化や高度化に留まらず、研究開発、マーケティング、戦略策定などの多岐にわたる分野に大きく貢献する可能性があります。この幅広い応用可能性について、具体的に説明します。

#### 1. 特許実務での効率化と高度化

- **権利期間の予測**や**特許のスコアリング**により、特許実務における重要特許の管理が容易になります。たとえば、権利期間が長期化する可能性がある特許を特定することで、企業は有望な特許を重点的に維持する戦略を採用でき、特許管理コストを抑えることが可能です。
- また、**引用情報を用いた重要技術の特定**により、知財部門が特許出願の優先順位や戦略を合理的に決定でき、出願戦略の精緻化や費用対効果の向上に貢献します。

#### 2. 研究開発 (R&D) での応用

- 研究開発分野では、特許情報の分析から得られる知見が技術選定や開発テーマの設定に役立ちます。特許データベースを LLM で分析し、**異分野技術の融合可能性を評価**することで、新たな研究テーマの発見や、異なる技術分野の組み合わせによる

革新が期待されます。

- さらに、LLM が知識グラフを用いて技術情報を体系的に整理することで、研究者は関連する過去の技術や特許を容易に把握し、重複研究を防ぎつつ、効率的な研究の方向性を見出すことが可能です。

### 3. マーケティング分野での活用

- 特許情報から**市場での需要が高い技術**や、他企業が採用している技術の傾向を分析することで、マーケティング部門は市場ニーズに合わせた製品企画やプロモーション戦略を策定できます。
- 例えば、成長性スコアを用いて市場での成長が見込まれる技術分野を特定し、企業のマーケティング戦略に反映させることで、消費者に響く製品の特徴を押さえた広告やプロモーションが可能になります。

### 4. 事業戦略や経営戦略の策定

- 特許情報の分析から得られる技術動向や他社の技術戦略に関する知見は、経営戦略や事業戦略の策定にも応用可能です。特に、特許の権利期間や引用ネットワークの成長性を基に重要技術を特定することで、**今後注力すべき技術分野や投資分野の判断**ができます。
- この技術評価を元に、経営層は M&A のターゲット選定や事業拡大計画の策定を支援し、将来にわたって競争力を強化するための意思決定を行うことができます。

### 5. 技術創造とイノベーション促進

- LLM による技術創造は、課題と解決手段の組み合わせを推論することで、新しい技術的解決策を提案することが可能です。これにより、技術の組み合わせから生まれるイノベーションが促進され、企業は**新製品開発や新しい技術領域への進出**を加速できます。
- 例えば、異なる技術領域からの課題解決策を組み合わせることで、新規性の高い技術が誕生しやすくなり、製品の差別化や競争優位性の確保に貢献します。

### 6. 知識伝達や教育における応用

- 知識グラフを活用したデータセット整備により、技術情報が整理されているため、特許情報が教育や社内研修に活用しやすくなります。企業内での知識共有や技術研修で、特許情報の分析結果やスコアリングを活用し、従業員が技術動向を学び、実践的な知識を得る機会を提供できます。
- また、学术界においても、特許データを用いた実践的なデータ解析や技術評価の教育が行われやすくなり、次世代の技術者や研究者の育成に役立つと考えられます。

### 7. 国際的な競争力の強化

- 官民学連携による特許データセットの整備と活用は、日本の技術解析や特許解析分野での競争力向上にも寄与します。特に、国内外の特許情報の網羅的なデータベースとそれを活用した技術創造 AI が整備されれば、特許解析や技術評価での国際的

な競争力が強化されることが期待されます。

- これにより、日本企業が国際市場での技術的優位性を確保し、より積極的な技術輸出や国際的なイノベーション活動に貢献する基盤が整います。

#### まとめ

野中准教授の研究は、特許実務の効率化をはじめとし、研究開発、マーケティング、戦略策定、教育など多岐にわたる分野での活用が期待されています。特に、特許情報を活用した技術創造 AI の整備により、革新を生み出す技術的なアイデアが支援され、企業全体や社会全体にわたる価値創出の可能性が広がります。