

# 第8回特許情報シンポジウム

論文資料集

2024年11月7日(木)

オンライン開催

一般社団法人アジア太平洋機械翻訳協会

一般財団法人日本特許情報機構

# 特許庁におけるA I 活用の取組について

---

2024年11月7日  
特許庁総務部総務課



# A I 技術活用の目的とこれまでの取組

## 目的

産業財産権を取り巻く環境の多様化・複雑化を踏まえ、特許行政事務の高度化・効率化を図り、ユーザー向けのサービス向上につなげる

## 取組

平成26年度：機械翻訳においてA I 技術の適用を試行開始

平成28年度：特許庁業務全体についてA I 技術の適用可能性を検討

平成29年度：人工知能関連技術活用可能性検証プロジェクトチーム（A I 検証P T）設置  
実証事業の開始、アクション・プランの策定・公表

令和元年度：アクション・プランの改定・公表

令和2年度：アジャイル型内製開発チームの立上げ、アクション・プランの改定・公表

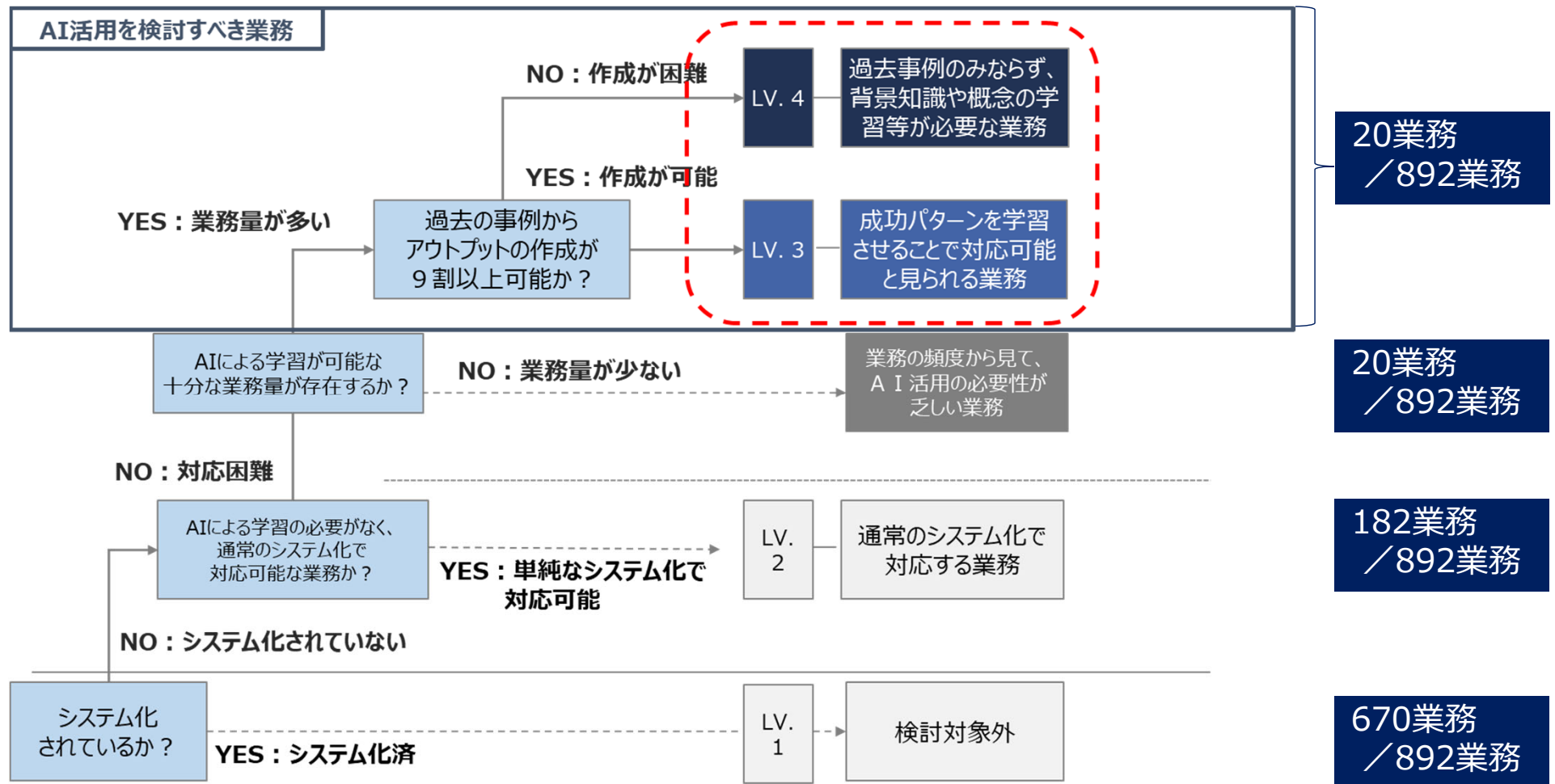
令和4年度：新たなアクション・プラン（令和4～8年度版）の策定・公表

令和6年度：アクション・プラン（令和6年度改訂版）の改定・公表

一部の業務については、A I 技術活用に関する技術実証を終え、業務内容に精通した職員とプログラマーのチームでアジャイル内製開発を実施

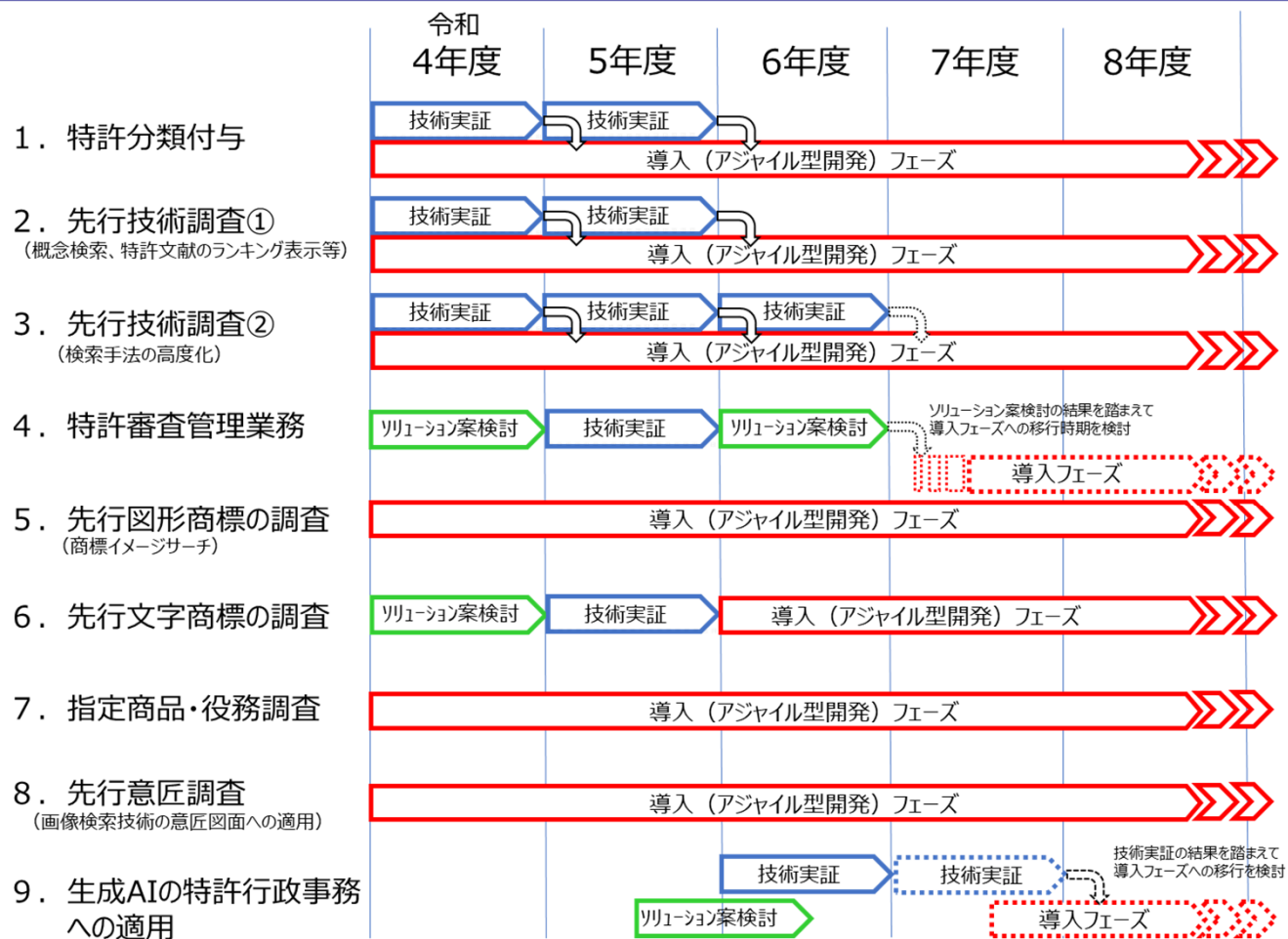
# A I 技術活用に向けた取組（業務の選定）

特許行政事務に係る業務全体を分析し、A I 技術の活用可能性のある20業務を選定



# アクション・プラン（令和6年度改定版）

令和6年5月に改定したアクション・プランに沿って、各事業の技術実証やツールの開発・試験的な導入を実施



(※) 各事業の取組は大まかな想定であり、開発の進捗状況や予算の状況、その他の諸情勢により、変更がありうる。

- 特許庁内で、解決すべき業務課題を検討・抽出
- 業務課題に対して、AI技術の実現可能性の等、技術実証を外注
- 技術実証の結果を考慮して、特許庁内でツールを内製・アジャイル開発

## 具体的な事例紹介

# 実体審査（特許）とA I 技術の関係

分類付与

本願発明の理解

先行技術調査

先行技術文献検索

スクリーニング

特許性の判断

## 特許文献への分類付与

### 背景

国内特許文献：人手で付与  
外国特許文献：日本の特許分類が付与されず

### 検証中の取組

A I 技術を活用して、国内特許文献に付与済みの特許分類を学習し、外国文献に特許分類を機械付与

### 効果

外国特許文献も日本の特許分類で先行技術文献調査が可能になり、審査の効率性が向上

## 先行技術文献検索

### 背景

キーワードを使って先行技術文献検索する際は、類義語を網羅的に含める必要がある

### 検証中の取組

A I 技術を活用して、検索に有用なワードを推定して、提示

### 効果

特許審査の質が向上

## スクリーニング

### 背景

先行技術文献調査で抽出した文献集合から、審査対象と技術的に近い文献を目視で探す必要がある（1000文献以上照会することも多い）

### 検証中の取組

- ・ A I 技術を活用して、審査対象と似ている順に文献集合を並べ替え
- ・ A I 技術を活用して、指定した図面と似ている先行技術文献の図面を表示

### 効果

先行技術調査の効率化、質の向上

# 実体審査（特許）とA I 技術の関係

分類付与

本願発明の理解

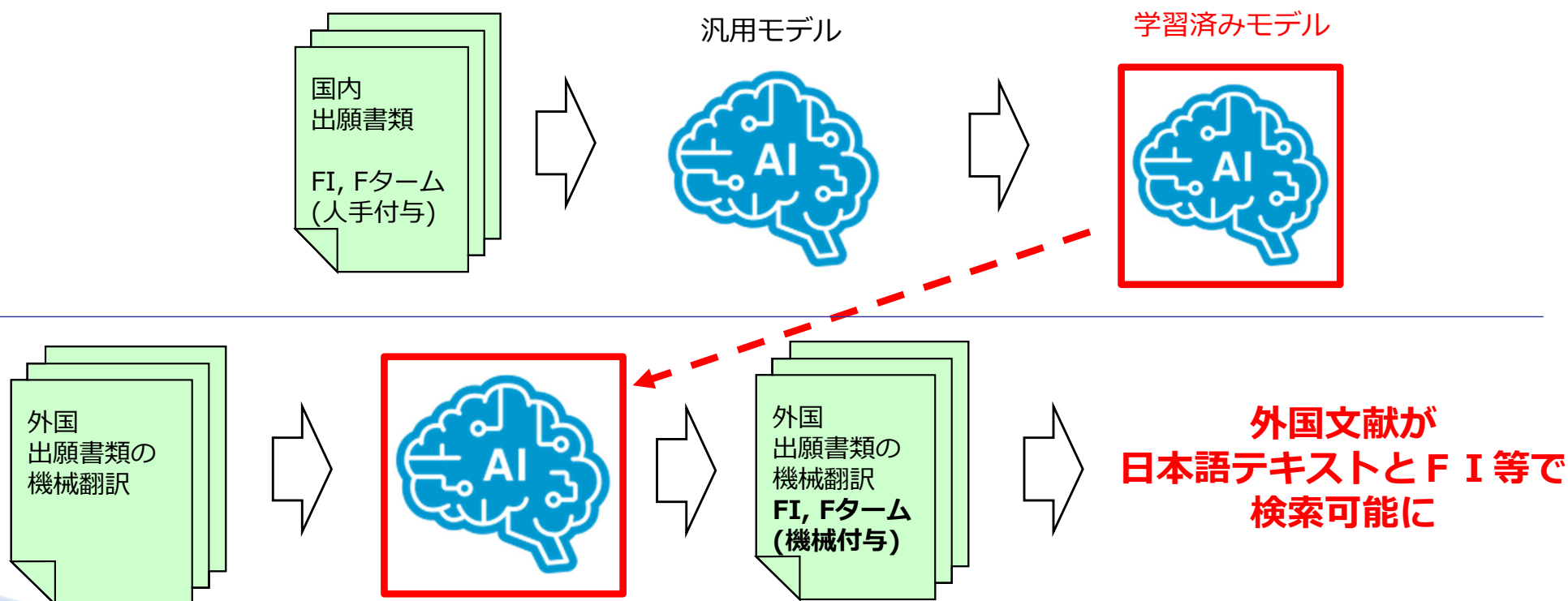
先行技術調査  
先行技術文献検索

スクリーニング

特許性の判断

## 特許分類付与支援

- 外国語の特許文献（約3,500万件）に対して、日本語に機械翻訳したテキストを入力として、分類情報を推定。推定結果を外国語の特許文献に対して紐付けることで、F I、Fタームおよび日本語テキストを用いた外国特許文献検索を可能に
- P C T等、日本の分類を審査官が付与する文献について、当該文献のテキストを入力として、分類情報を推定し、審査官の付与を支援





# 実体審査（特許）とA I 技術の関係

分類付与

本願発明の理解

先行技術調査  
先行技術文献検索

スクリーニング

特許性の判断

## 検索式作成支援

適切な検索式を立てるには、適切な特許分類や類義語を選択する必要があり、習熟が必要  
⇒審査対象案件の「特許請求の範囲」および「明細書」の記載を元に、先行技術調査で用いる検索式作成に有用なキーワード等を推定し、審査官に提示

## 提案システム

習熟した知識及び経験の収集：

- ・分類に関する知識
- ・類義語のバリエーション

習熟した審査官



サーチ履歴等

### サーチクエリの拡大

分類

**A61B34/35**  
B25J13/00  
A61B17/00

キーワード

**robot,**  
**manipulator,**  
arm, master,  
operator, remote

案件情報

審査官補／異動審査官

特許出願



検索式提案

(A61B34/35+B25J3/00)  
AND  
(robot+manipulator+arm+  
master)

# 実体審査（特許）とAI技術の関係

分類付与

本願発明の理解

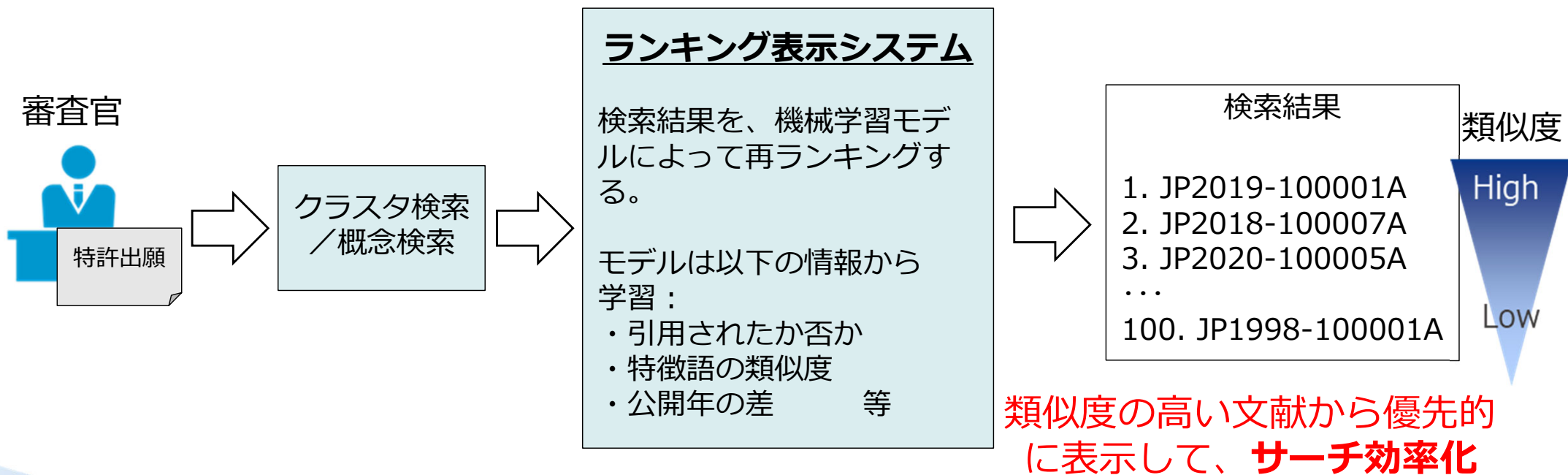
先行技術調査  
先行技術文献検索

スクリーニング

特許性の判断

## 特許文献のランキング表示

調査対象となる特許文献に含まれる情報（書誌情報、分類、請求項、明細書、引用回数等）を用いて、審査対象案件との類似度に関するスコアを機械学習モデルによって付与し、スコアの高い順に表示



# 実体審査（特許）とAI技術の関係

分類付与

本願発明の理解

先行技術調査  
先行技術文献検索

スクリーニング

特許性の判断

## 画像検索技術の特許図面への適用

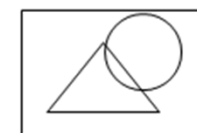
- 審査対象案件の中で指定された「図面」を元に、関連性の高い図面を優先的に提示
- また、事前に各図面に回路図等のカテゴリーを機械付与しておき、図面検索の際に指定されたカテゴリーの図面を優先的に提示

### 図面検索システム

選択した図面に類似する順に、対象図面群を並び替えて表示

以下の情報に基づき、図面を抽出：  
・ 図面に埋め込んだキーワード  
・ 図面の類似度

### 検索結果



JP2005-XXXXXX



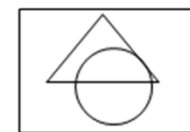
JP2010-XXXXXX



JP2009-XXXXXX



JP2005-YYYYYY



JP2010-YYYYYY

...

審査官

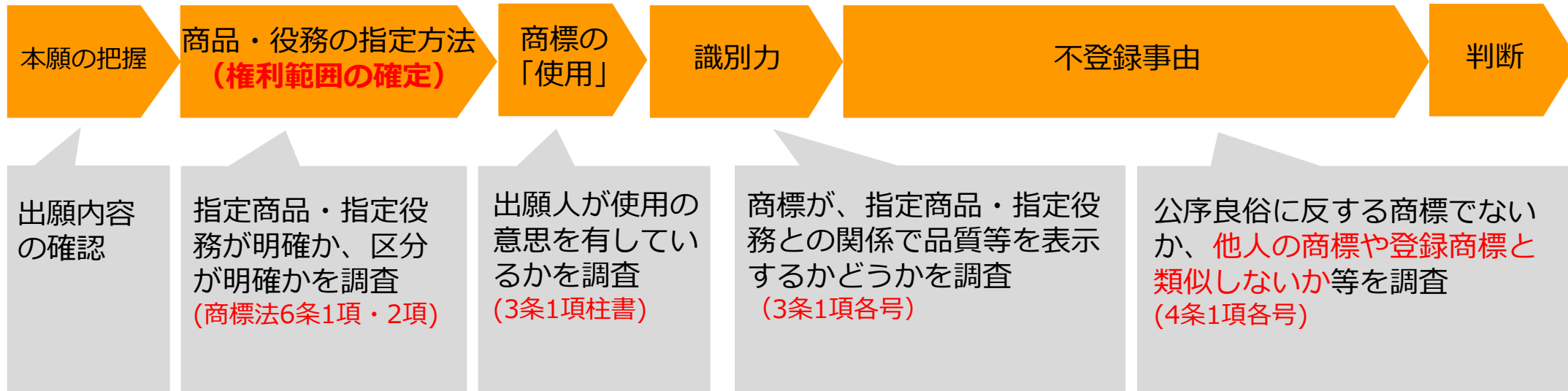


特許出願



関連性の高い図面から優先的に表示して、**サーチ効率化**

# 実体審査（商標）とA I 技術の関係



# 実体審査（商標）とAI技術の関係



## 指定商品・役務名の調査

### 業務内容

庁内DB収録の商品・役務を基に、指定商品・指定役務に「類似群コード」を付与

### 従来

1件ずつ手作業で付与（100件以上の指定商品・指定役務が存在することも）

### 効果

- ・一部の出願については商品・役務に係る審査時間を3割以上短縮
- ・テレワーク時における審査処理を促進

## 先行図形商標の調査

### 業務内容

先行する他人の図形商標を検索し、類否を判断

### 従来

- ・図形分類をキーに検索式を作成し、スクリーニングしていたが、対象が十分に絞り込めない（3000件スクリーニングすることも）
- ・異なる図形分類が付与されていた場合にはヒットしない

### 効果

- ・イメージサーチを利用した検索システムにより酷似する図形を短時間で検出
- ・図形分類に依拠せずに候補を検出することで審査の品質向上に寄与

### 課題

検索対象側が結合商標の場合、商標の要部を捉えた検索が難しい  
※文字商標の検索は、従来型の検索システムを利用

# 実体審査（商標）とAI技術の関係

本願の把握

商品・役務の指定方法

商標の「使用」

識別力

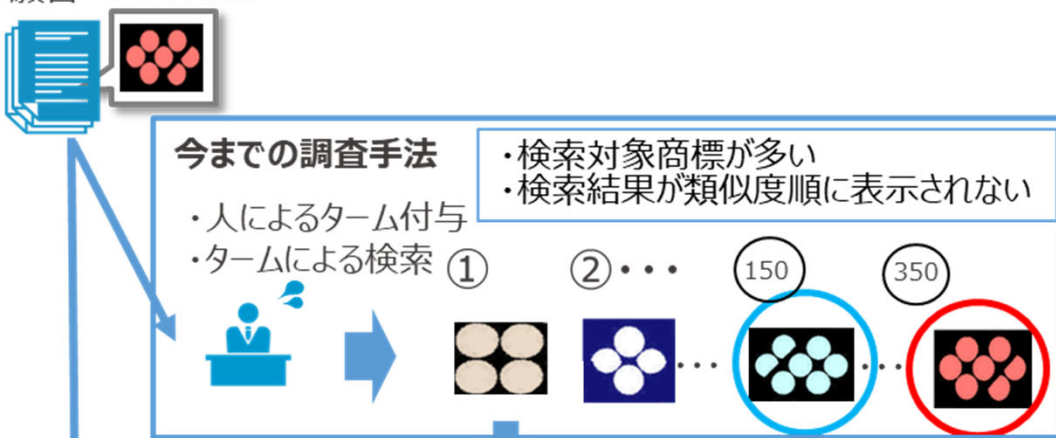
不登録事由

判断

- 平成29年度及び平成30年度に、「先行図形商標の調査」及び「指定商品・役務調査」について実証研究を実施
- 成果物は、アジャイル型開発手法により試行的に審査支援ツールとして導入し、効果を検証中
- 商標審査において活用することで、審査の品質向上を図る

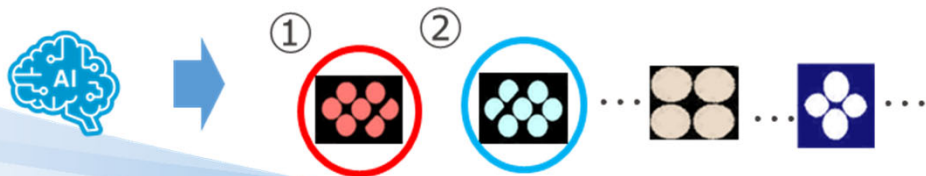
## 先行図形商標の調査

願書 図形商標



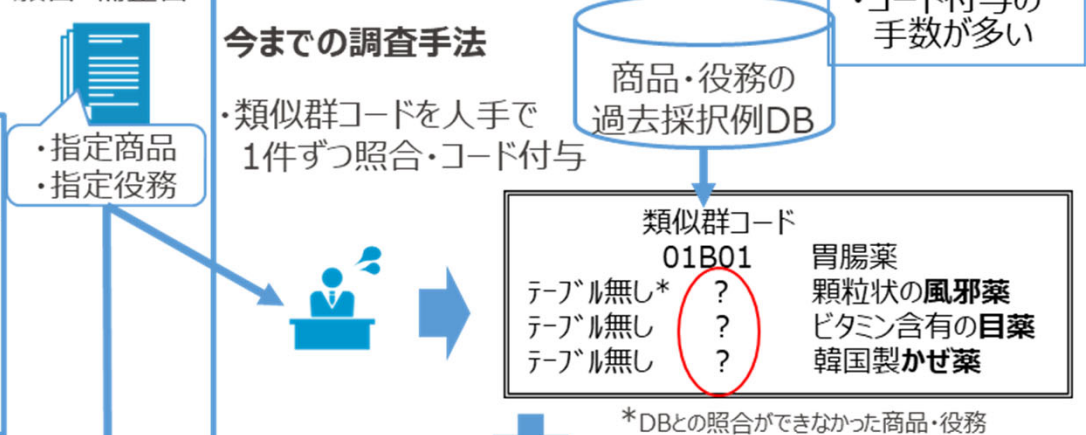
AIによる業務サポート

- ・類似度順に検索結果を表示
- ・見過ごし防止等による審査品質向上に貢献



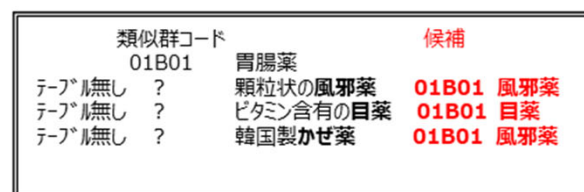
## 指定商品・役務調査

願書・補正書



AIによる業務サポート

- ・類似群の候補をAIツールが提示
- ・コード付与の効率化・品質向上に貢献



# 実体審査（商標）とAI技術の関係

本願の把握

商品・役務の指定方法

商標の  
「使用」

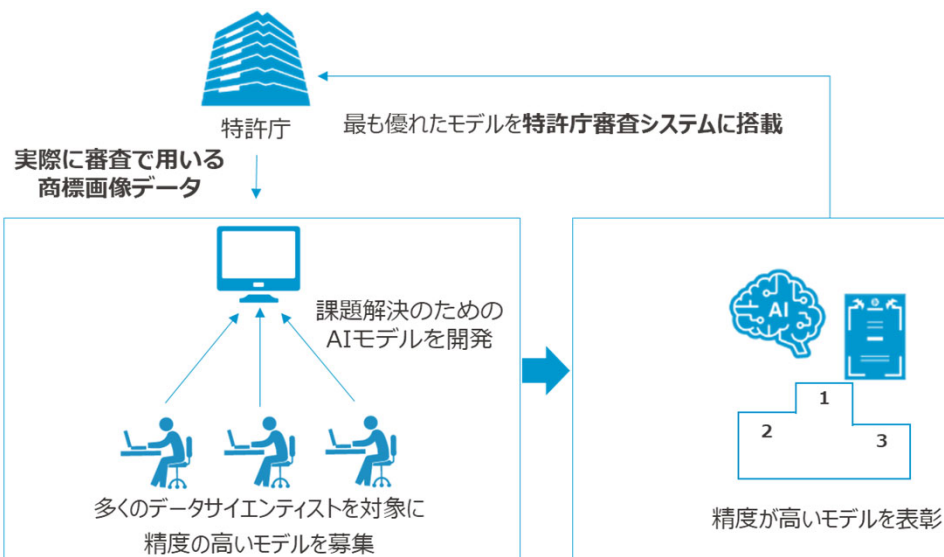
識別力

不登録事由

判断

- AIを利用したイメージサーチツールの精度向上のため、特許庁初の機械学習コンペティション（※）「AI×商標：イメージサーチコンペティション」を開催（令和3年11月26日～令和4年1月31日）
- 入賞者のAIモデルを実装することにより、現行ツールの2倍程度の精度向上が期待
- 審査支援ツールのAIモデル検討におけるコンペ方式の有効性を確認

## AIコンペ概念図



## AIコンペ参加者・投稿者数

- 参加者数：**637名**
- 投稿件数：**1,453件**

### ※機械学習コンペティションとは

特定の課題及び関連データを公開し、その課題解決のためのモデル開発を広く一般から募り、優秀なものを採用する開発手法。優れた技術やアイデアを有する人材を新たに発掘し、システム開発に参画する機会を創出することで、**効率的なシステム開発及びイノベーションの促進**も期待。**実装まで含めたコンペ開催は極めて珍しい取組。**

# 実体審査（意匠）とA I 技術の関係

分類付与

本願意匠の理解

先行意匠調査

登録性の判断

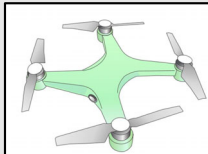
## 画像検索技術の意匠図面への適用

- 先行意匠調査の効率化・品質向上を図ることを目的とし、先行意匠の調査において、類似画像検索技術を活用
- 審査対象案件の中で指定された「図面」を元に、関連性の高い図面を優先的に表示
- 教師データの質と量を改善し、精度向上を目指している（開発中）

審査官



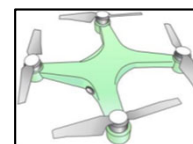
意匠出願



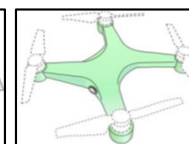
## 類似画像検索システム

選択した図面に類似する順に、  
対象図面群を並び替えて表示

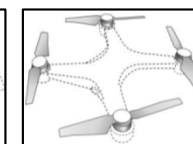
## 検索結果



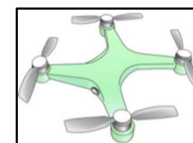
D1AAAAA



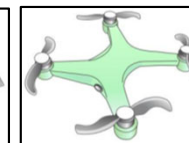
D1BBBBB



D1CCCCC



D1DDDDD



D1EEEEEE

• •  
•

関連性の高い図面から優先的に表示して、サーチ効率化



ありがとうございました

---

特許庁総務部総務課



# 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

2024/11/7

愛知工業大学 野中尋史

# 自己紹介

野中 尋史:hnonaka@aitech.ac.jp

愛知工業大学 経営学部 准教授

- 2015年 長岡技術科学大学 工学研究科 講師
- 2018年 長岡技術科学大学 工学研究科 准教授
- 2022年 愛知工業大学 経営学部 准教授  
(2006-2011年 豊橋技術科学大学知財連携コーディネーター)

## 委員歴

- 2020年4月 - 現在妙高市, 連携研究員
- 2021年4月 - 2022年3月電子情報通信学会, 会誌編集委員
- 2021年4月 - 2022年3月電子情報通信学会, 代議員
- 2019年4月 - 2022年3月長岡AIイノベーションハブ, 代表
- 2019年6月 - 2021年6月電子情報通信学会, 信越支部庶務幹事

# 特許・論文



# 異分野横断 技術創造手法

重要  
技術  
特定



機械工学技術

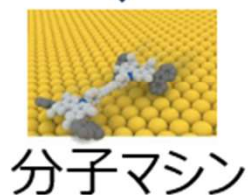
機械工学  
技術



融合



化学  
技術

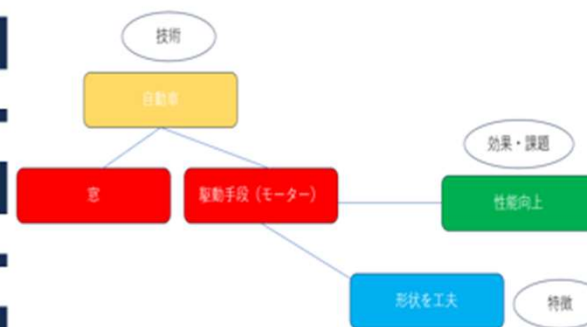


分子マシン

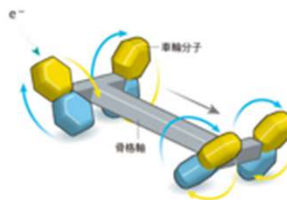
技術融合可能な  
技術群特定

# 創造された技術構成

## 知識グラフ



## 図面



LLMでの技術創造・推論  
(技術構成の具体化)

# 研究目的

- 技術を創造するAIを開発しよう！
  - 研究開発
  - 知財戦略
  - マーケティング
  - ...
- 大規模言語モデルだけでは不安・・・
- 特許スコアリングモデルと大規模言語モデルを組み合わせよう！

# 特許・論文



## 異分野横断 技術創造手法

重要  
技術  
特定



機械工学技術

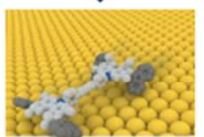
機械工学  
技術



化学  
技術



融合



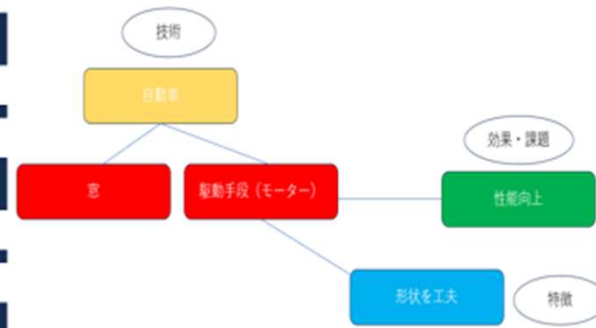
分子マシン

技術融合可能な  
技術群特定

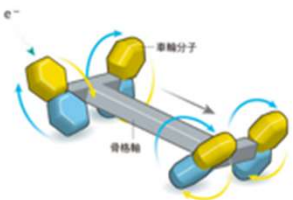
LLMでの技術創造・推論  
(技術構成の具体化)

## 創造された技術構成

### 知識グラフ



### 図面



# 重要技術の特定とスコアリング

- 異分野にも波及する重要技術の定量評価の研究が盛ん
- 特許重要性に関するスコアリングの観点は複数ある
  - 権利期間
  - 引用情報
  - 特許権に関連するアクション
  - 特許文書の品質
  - ...

# 重要技術の特定とスコアリング

- 異分野にも波及する重要技術の定量評価の研究が盛ん
- 特許重要性に関するスコアリングの観点は複数ある
  - 権利期間
  - 引用情報



# 重要技術の特定とスコアリング

- 異分野にも波及する重要技術の定量評価の研究が盛ん
- 特許重要性に関するスコアリングの観点は複数ある
  - 権利期間：長期での予測が難しい
  - 引用情報



Technological Forecasting and Social Change

Volume 203, June 2024, 123390



## A study on patent term prediction by survival time analysis using neural hazard model

Koji Marusaki <sup>a</sup>  , Kensei Nakai <sup>b</sup>, Shotaro Kataoka <sup>c,f</sup>, Seiya Kawano <sup>d</sup>, Asahi Hentona <sup>b</sup>, Takeshi Sakumoto <sup>c</sup>, Yuta Yamamoto <sup>b</sup>, Kaede Mori <sup>g</sup>, Hirofumi Nonaka <sup>e,f</sup>

Show more 

 Add to Mendeley  Share  Cite

<https://doi.org/10.1016/j.techfore.2024.123390>

[Get rights and content](#) 

# 背景と目的

- 特許情報を用いた技術の価値評価は企業での意思決定に有益[1]
  - 意思決定：企業の研究開発やM&A等
  - しかし、分析には費用・労力がかかる上、定量的な評価も難しい
    - 容易に特許技術を評価できる手法が求められる
- 特許権の生存分析による、各特許の価値評価
  - 日本の特許権は、出願日から最長20年間（一部例外あり）
    - 権利期間 = 権利者が「**特許料**」を何年間納付し続けるか
      - 価値のない特許にはお金を払わない
      - 仮説：権利期間の**長い特許 = 価値の高い特許**（権利）
  - 目的：特許権利期間の背景にある特許素性を用いた権利期間予測

# 先行研究

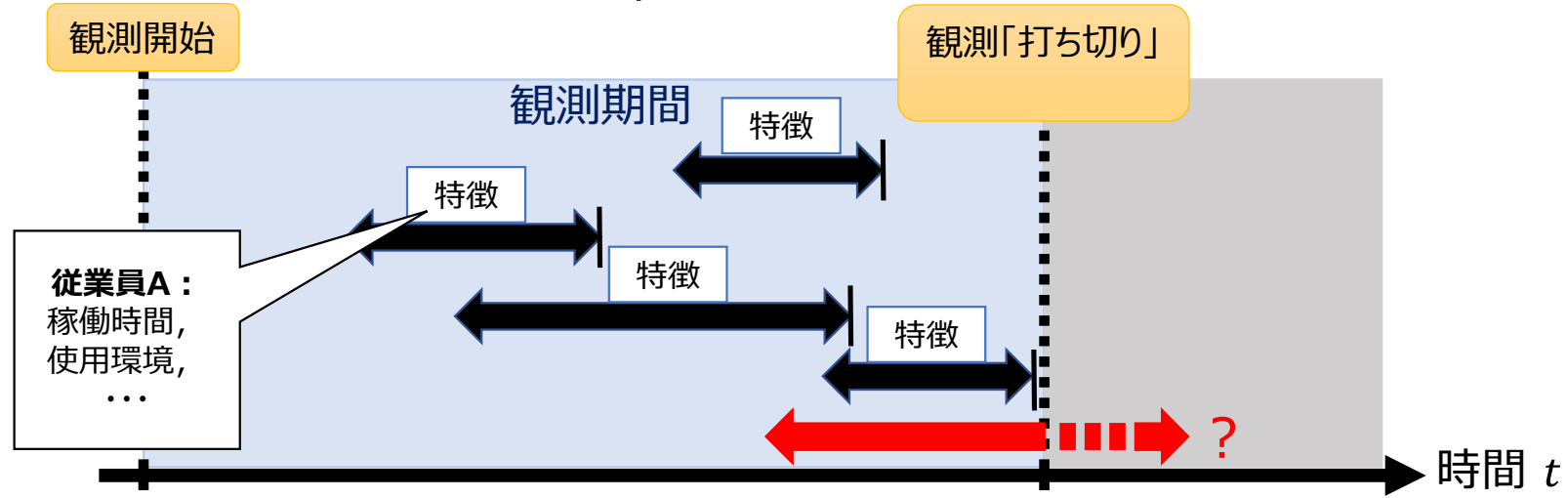
- 「**特許権利期間の長さが経済価値の指標となる性質**」
  - Bessen [2] : 予め算出した特許使用料収益推移を用いて、権利期間を推定
    - 権利存続中の特許から、正確な特許収入・特許権利期間の予測は困難
- **特許文書の素性から特許権利期間を生存分析**
  - 和田 [3] : 特許の引用情報に着目して、特許権利期間への影響を分析
  - Zeebroeck [4] : 欧州における特許権利期間を決定する要因を総合的に分析
- 2010年代より、ニューラルハザードモデルが提唱・適用[5,6]
  - 深層学習を生存分析に取り入れた手法
  - 説明変数－目的変数間の関係式における、非線形な作用も分析に適用可能ニューラルハザードモデルであるDeepSurv [7]を用いて特許素性から権利期間を分析

# 生存分析とは

- ある時系列に対して,  
「イベントが発生する**タイミング**」(目的変数)と  
「そのイベントが発生するための**条件**」(説明変数)との関係进行分析
  - →未知のケースに対して「いつイベントが発生するか」を予測可能
- 「イベント」の定義：
  - 同じ時系列において、「1回のみ」発生
  - 繰り返し発生する場合, それぞれを別々の事例と捉える
- 「イベント」の例：
  - 機械の故障, 定期課金サービスの解約, 借金の完済 等

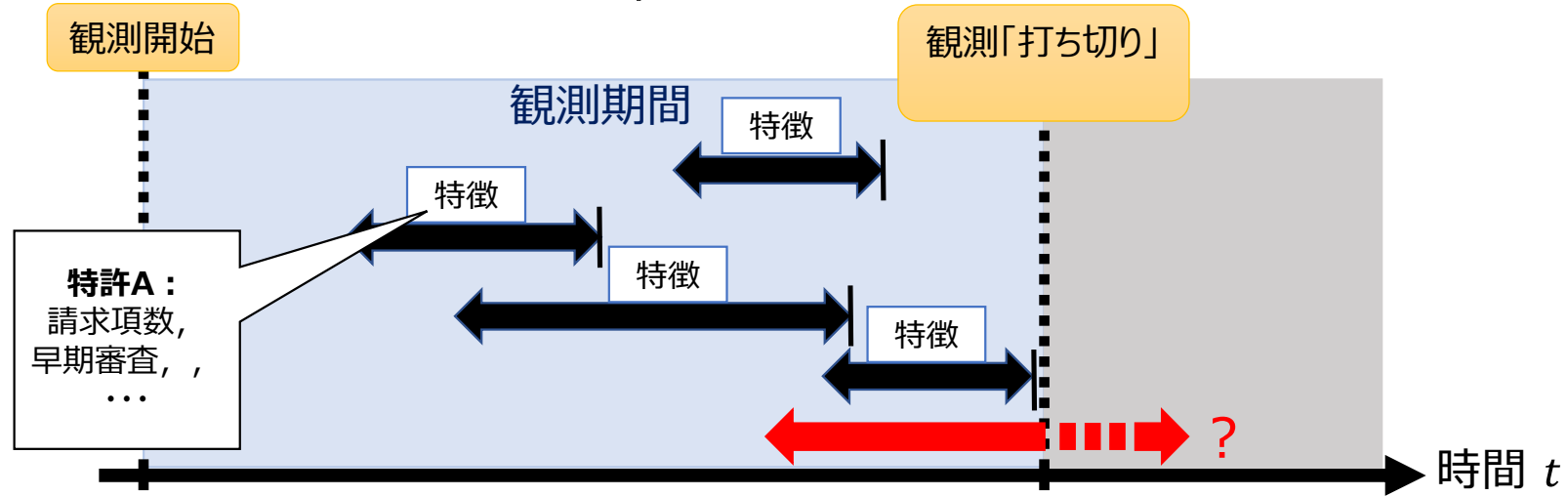
# 生存分析の例

- 「機械の故障時間予測」
  - イベント：機械の故障
  - 説明変数：機械の稼働時間，使用環境（標高，気温，湿度，・・・）
- 「説明変数がどのような場合に，故障する可能性が高いか」を分析



# 生存分析の例

- 「特許権利期間予測」
  - イベント：特許権の放棄（年金負支払）の判断
  - 説明変数：請求項数，早期審査の有無，
- 「説明変数がどのような場合に，権利を放棄する可能性が高いか」を分析



# 生存分析の概要

- 生存分析に必要なデータ :

- $x$ : ベースラインデータ (説明変数となる特徴ベクトル)
- $T$ : イベント発生時刻
- $E$ : イベント発生の有無 (0 or 1)

- 生存分析に必要な 2 つの関数 :

- 生存関数 Survival function :

- 時刻  $t$  の時点で生き延びる (イベントが発生しない) 確率

$$S(t) = \Pr[T > t]$$

- ハザード関数 Hazard function:

- $t$  時まで生存していた条件下で, 次の瞬間 ( $t + \Delta t$ ) に死亡 (イベントが発生) する確率

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{\Pr[t \leq T < (t + dt) | T \geq t]}{dt} = -\frac{d}{dt} \log[S(t)]$$

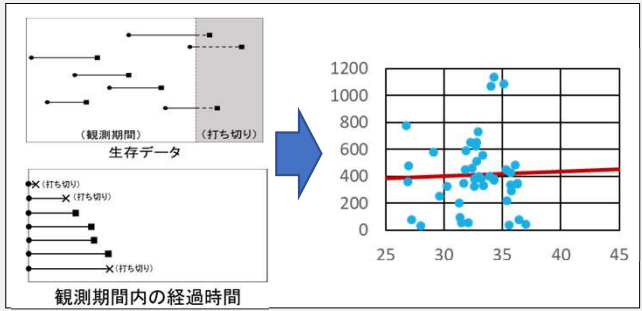
# 単純な回帰分析との違い

「打ち切り」: 生存データ特有の、「実験観測終了」によるイベント時系列の欠損

**× 単純な回帰分析**

- データの「打ち切り」を考慮しない
  - 「イベント発生」と「打ち切り」が区別できない

分析結果から、  
正確な生存時間の予測は難しい

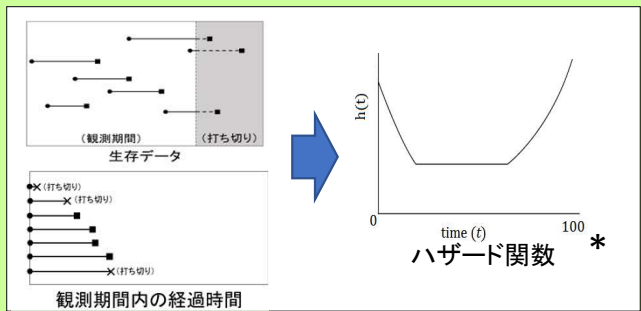


観測期間内の経過時間

**生存分析**

- データの「打ち切り」を考慮
  - 説明変数との関係を確率分布に反映

打ち切りがあっても、  
直前までの生存データを反映可能



観測期間内の経過時間

ハザード関数 \*

\*“Statistics How To” <<https://www.statisticshowto.datasciencecentral.com/hazard-function/>>より



# 分析手法：DeepSurv

Cox比例ハザードモデル[8]の深層学習版（多層パーセプトロン）  
従来手法との違い：「リスク関数 $h(x_i)$ および変数寄与度 $\theta$ の算定方法」

## Cox比例ハザードモデル（従来手法）

- リスク関数 $h(x_i)$ ：
  - 線形的に推定

$$\hat{h}_\theta(x_i) = \theta_1 x_{1,i} + \theta_2 x_{2,i} + \dots + \theta_n x_{n,i}$$

- 変数寄与度 $\theta$ ：
  - 対数尤度関数で最尤推定

$$l(\theta) := - \sum_{i:E_i=1} \left( \hat{h}_\theta(x_i) - \log \sum_{j \in \mathcal{R}(T_i)} \exp[\hat{h}_\theta(x_j)] \right)$$

## DeepSurv

- リスク関数 $h(x_i)$ ：
  - 変数寄与度 $\theta$ をネットワークの重みとして，損失関数で推定
- 変数寄与度は推定せず，直接リスク関数を算出



モデルのより柔軟な表現が期待できる

# 分析手法：DeepSurv

リスク関数  $\hat{h}_\theta(x_i)$  : 各特徴ベクトルを入力とし, ハザード関数  $\hat{h}_\theta(x_i)$ を推定して出力

$$\text{損失関数 } l(\theta) := - \sum_{i:E_i=1} \left( \hat{h}_\theta(x_i) - \log \sum_{j \in \mathfrak{R}(T_i)} \exp[\hat{h}_\theta(x_j)] \right)$$

$\hat{h}_\theta(x_i)$ : リスク関数  
 $\theta$ : ネットワーク (説明変数) の重み  
 $x$ : 説明変数  
 $\lambda_0(t)$ : ベースラインハザード関数  
 $\mathfrak{R}(T)$ : 時間  $T$  におけるリスク集合

ハザード関数  $\lambda(t|x)$  : 時間成分  $\lambda_0(t)$  と特徴成分が比例すると仮定

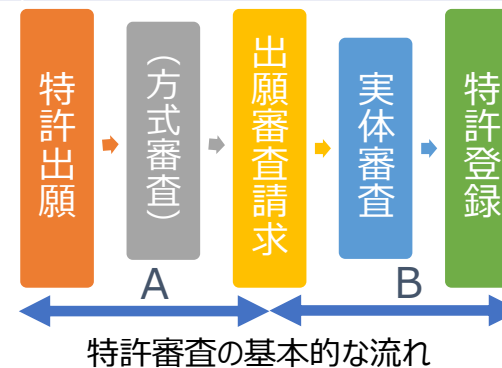
$$\lambda(t|x) = \underbrace{h_0(t)}_{\text{時間成分}} \cdot \underbrace{\exp[\hat{h}(x)]}_{\text{特徴成分 (リスク関数)}}$$

$\hat{h}(x)$ : リスク関数  
 $x$ : 説明変数  
 $h_0(t)$ : ベースラインハザード関数

# 入力する説明変数

Name	Content
Inventor	the number of inventors
Claim	the number of claims
IPC	(categorical variable) Whether each section of the IPC is assigned
Patent Owner Attributes	(categorical variable) Whether a patent owner belongs to "company" or "school" or "public_inst"
Citing	the number of examiner backward citations
Cited_std	the standardized value of the number of examiner forward citations
Reqstb_days	the term from the day of filing the examination request to the deadline of it (The earlier the request for examination, the higher the value)
Exam_days	the term from the day of requesting examination to the day of being granted

※審査請求期限：  
2001年9月30日出願分までは7年，それ以降は3年



# 実験

- Pythonライブラリ「PySurvival」[10]のDeepSurvモデルを用いて分析
- 使用データ：
  - 知的財産研究所「IIPパテントデータベース」(2017) より
    - 1980～1999年出願の特許, かつ2004年末時点で**特許期間が満了**しているもの
    - 使用する項目においてデータに欠損・例外のない**予測対象特許の内訳**  
(権利失効時期別)
    - データ数：210,752件
- 対象特許：
  - 観測期間：1980～1999年 (20年間)
  - 予測期間：1980～2004年 (25年間)
  - 特許分類：すべて

権利失効時期	特許件数 (件)
観測期間中 (打ち切りなし)	81,051
観測期間外 (打ち切りあり)	129,701
<b>合計</b>	<b>210,752</b>

# 実験

Method	CPH	Deepsurv
C-index (CI)	0.680 (0.681, 0.679)	0.677 (0.683, 0.670)
RMS E (CI)	862.062 (922.285, 801.839)	568.641 (664.013, 473.270)

# 重要技術の特定とスコアリング

- 異分野にも波及する重要技術の定量評価の研究が盛ん
- 特許重要性に関するスコアリングの観点は複数ある
  - 権利期間
  - 引用情報：引用数は非階層的で過去の特許を過大評価.

## Development of the patent values evaluation method considering growth of technical community

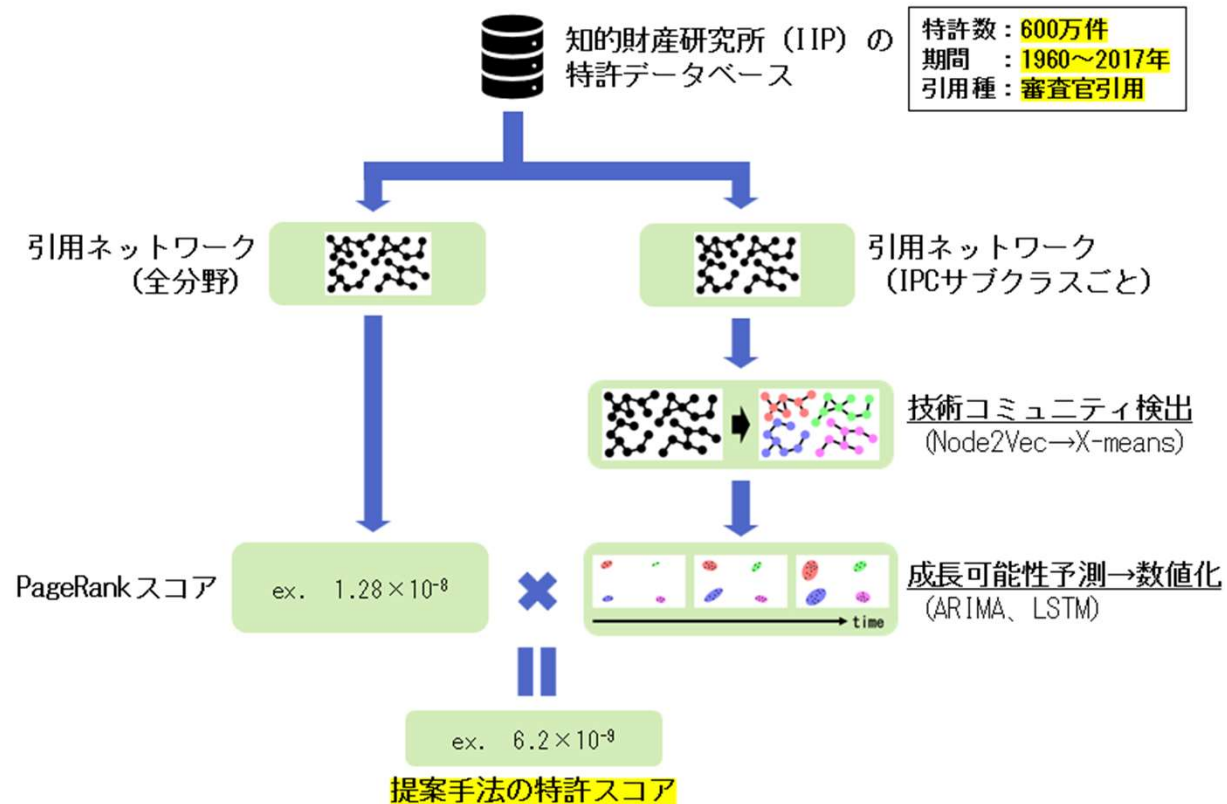
Publisher: **IEEE**

[Cite This](#)

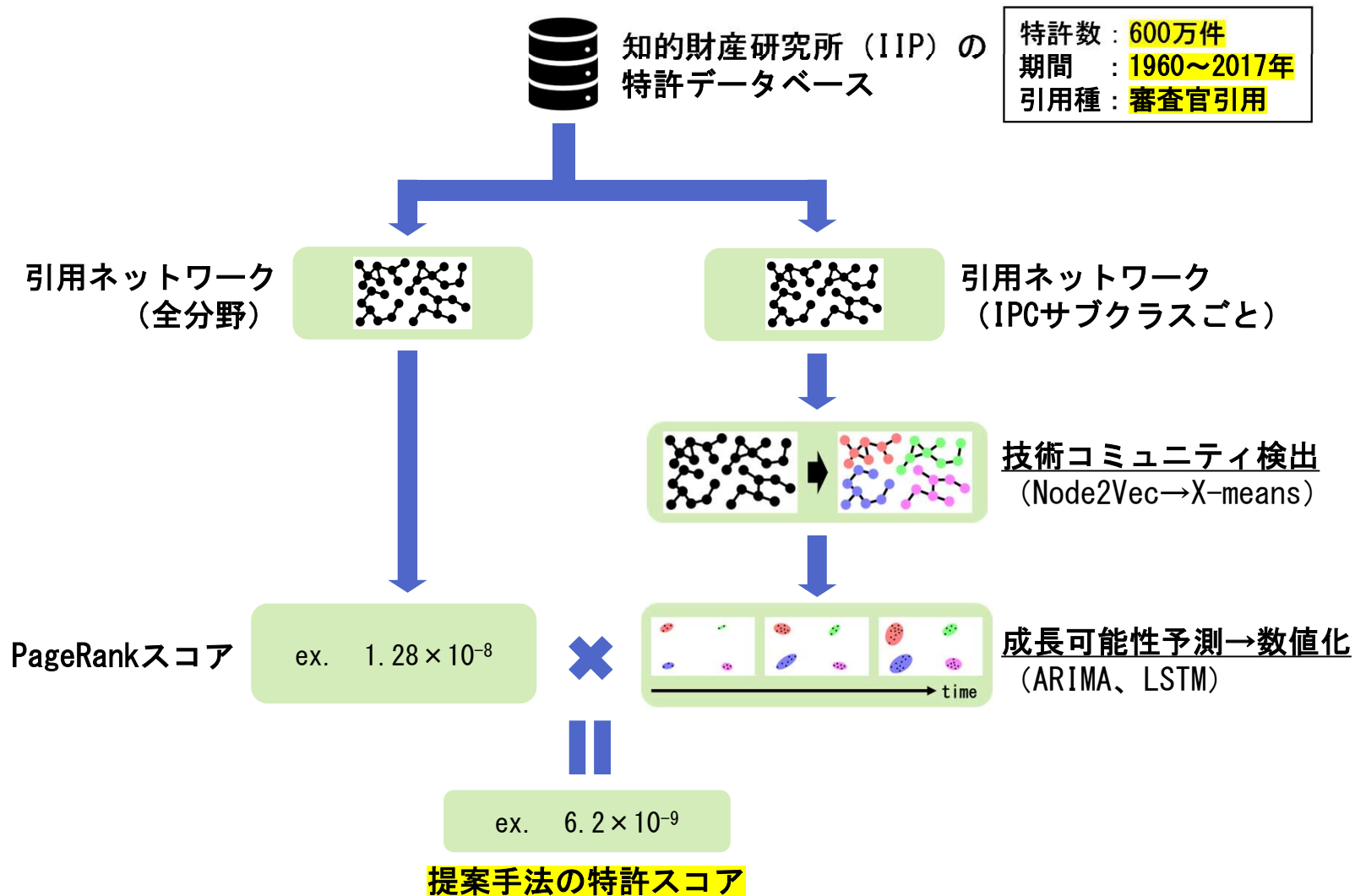
[PDF](#)

[Yuta Yamamoto](#) ; [Asahi Hentona](#) ; [Koji Marusaki](#) ; [Kohei Watabe](#) ; [Seiya Kawano](#) ; [Tokimasa Goto](#) **All Authors**

# 引用ネットワークの成長性も加味したスコアリング

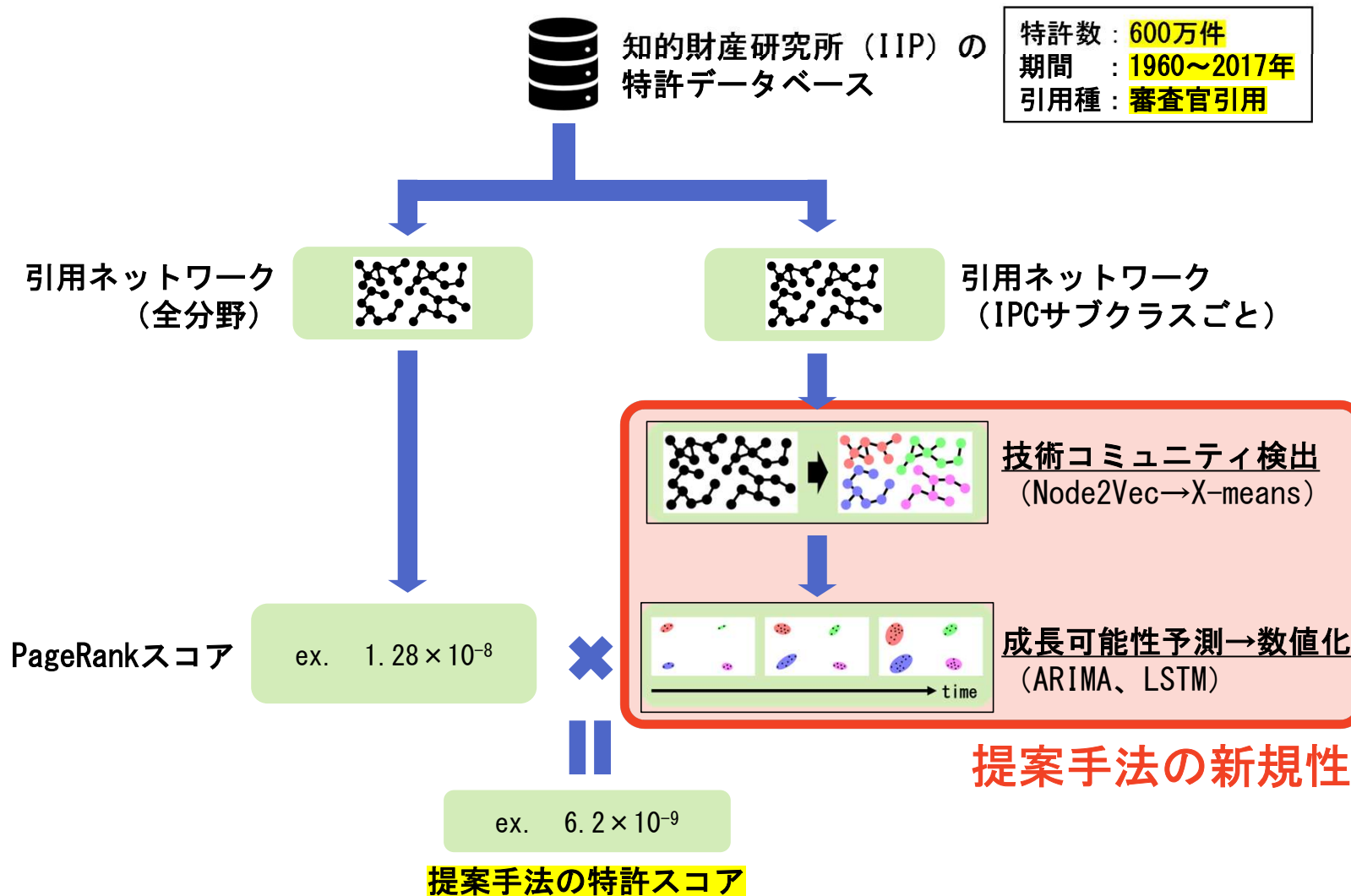


# 引用ネットワークの成長性も加味したスコアリング

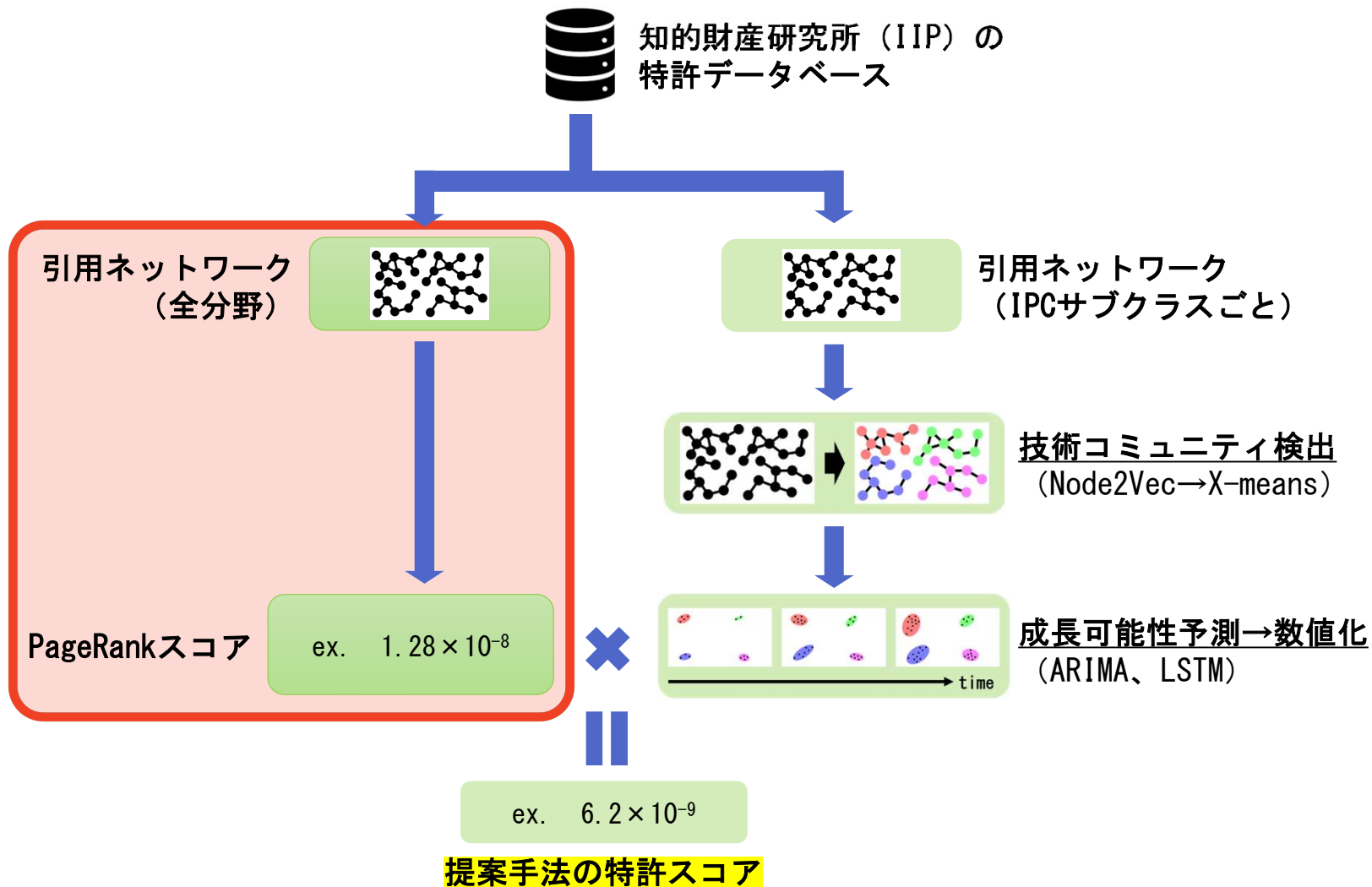




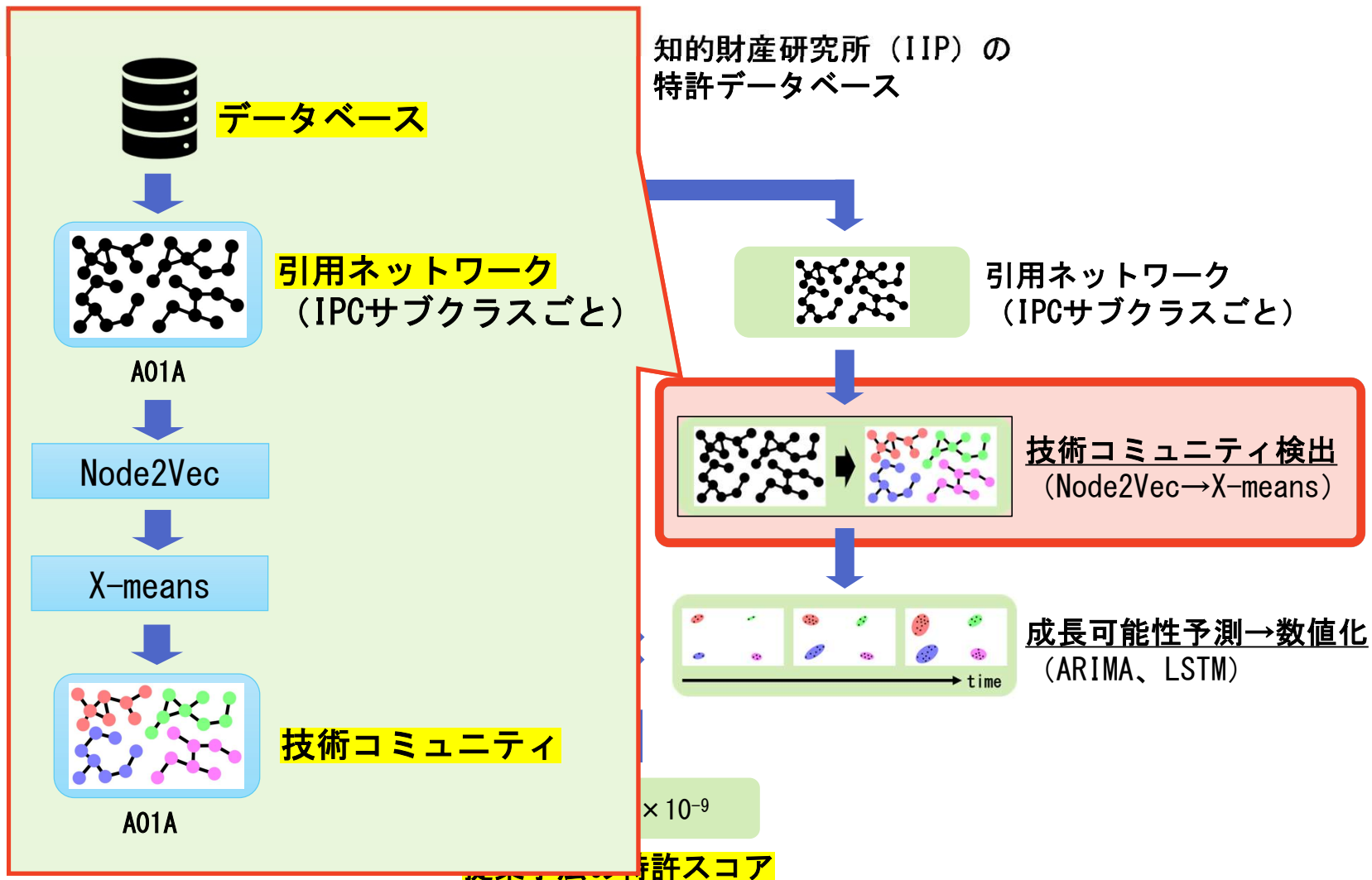
# 引用ネットワークの成長性も加味したスコアリング



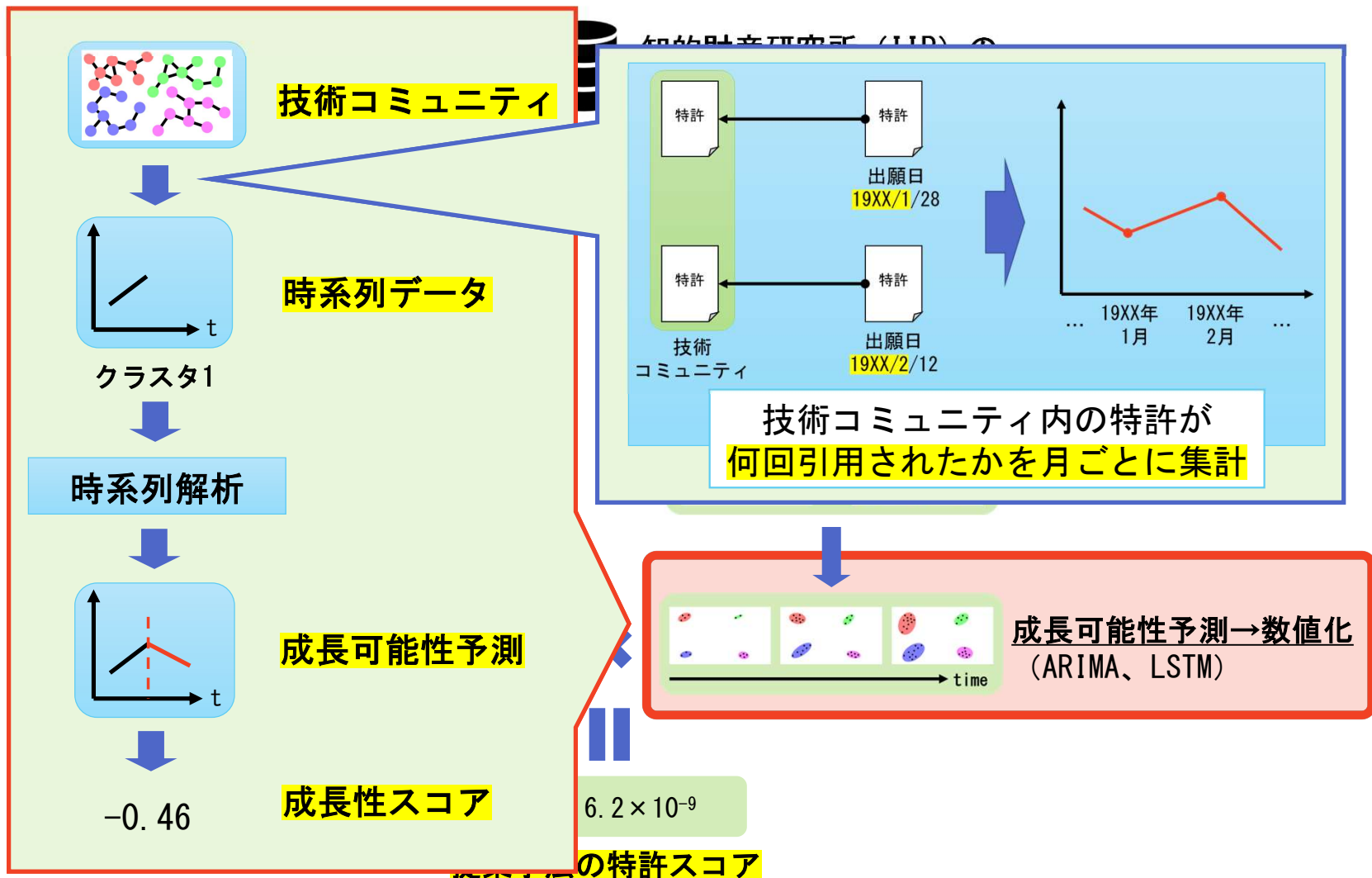
# 引用ネットワークの成長性も加味したスコアリング



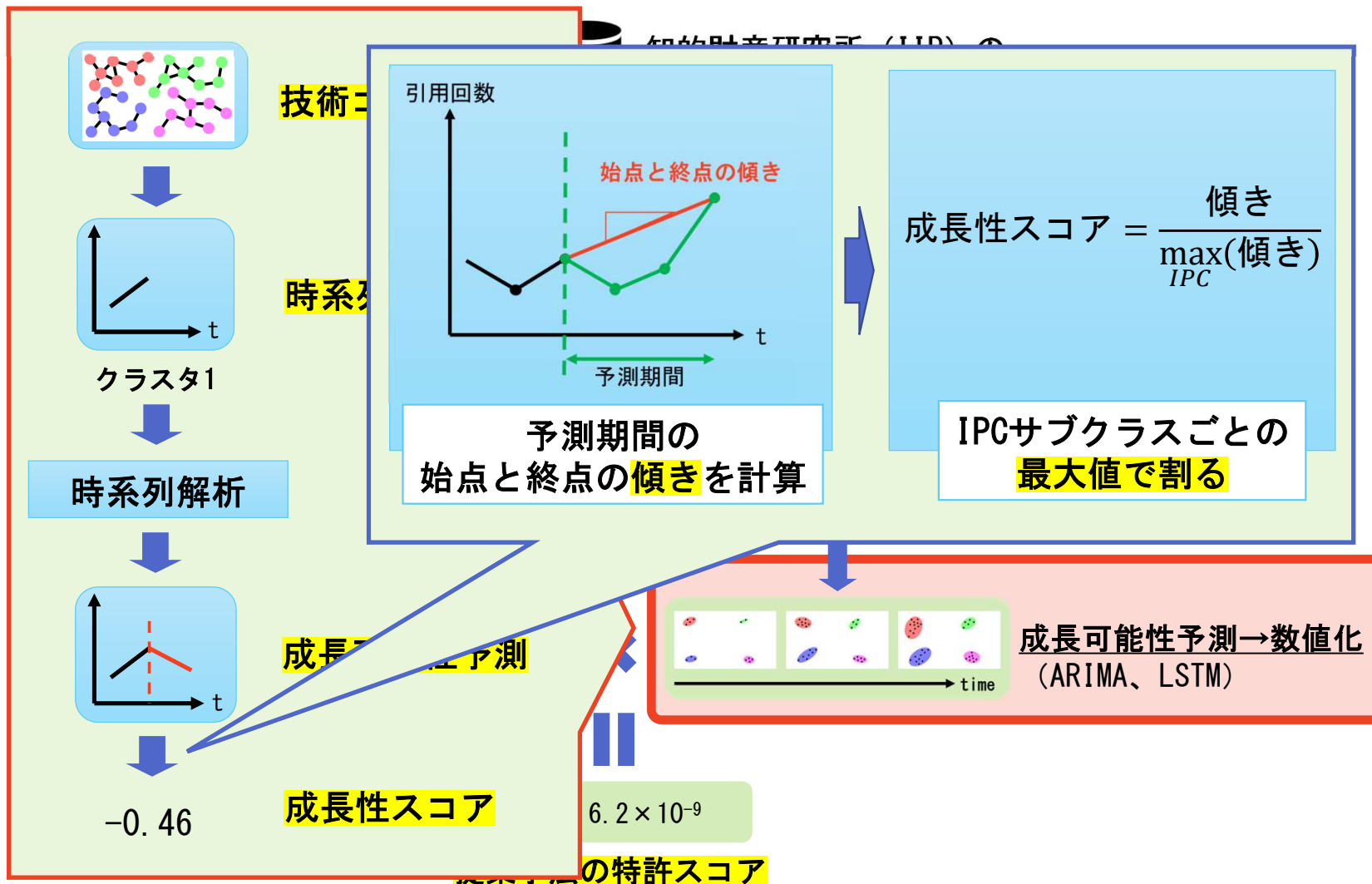
# 引用ネットワークの成長性も加味したスコアリング



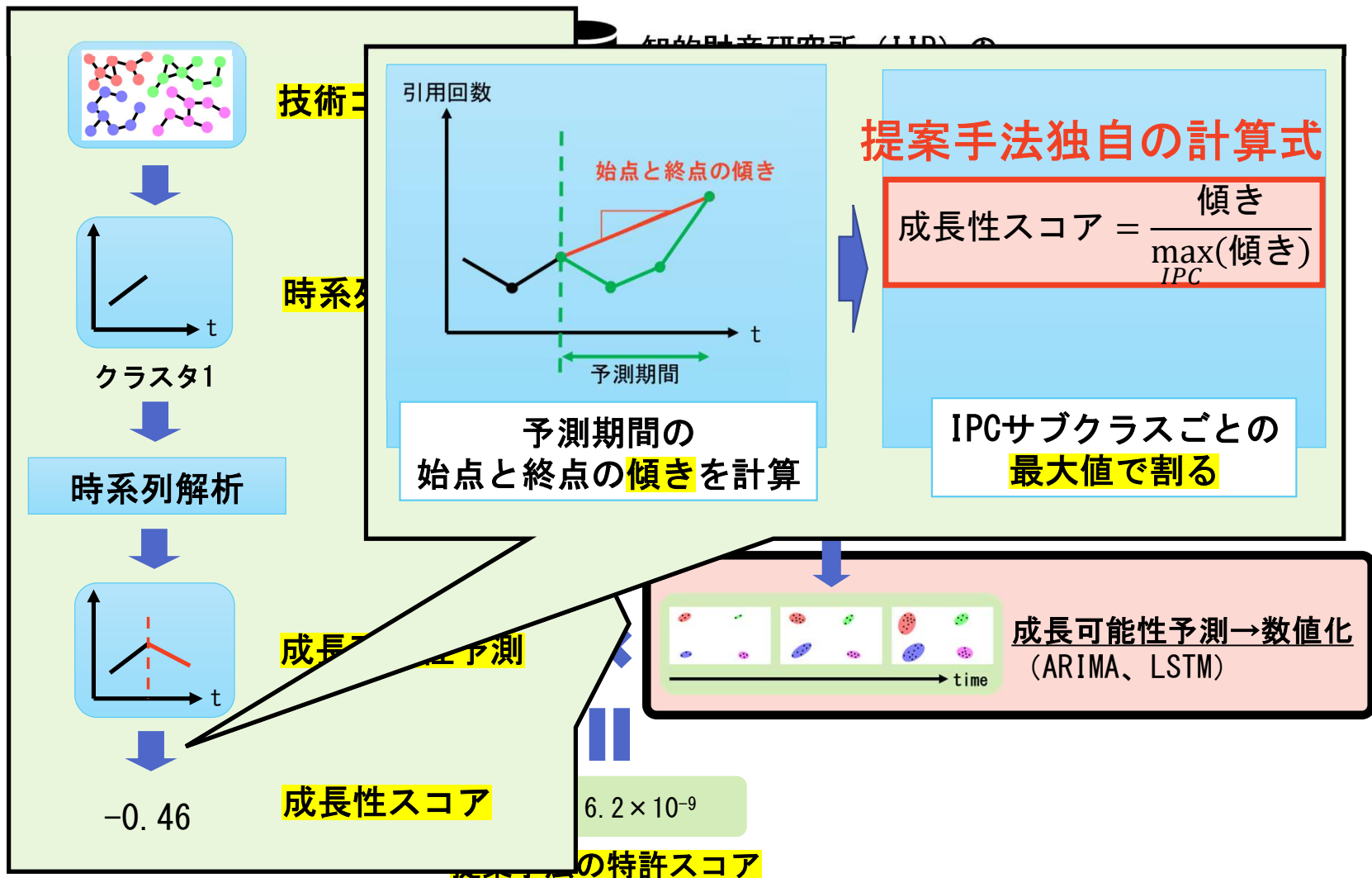
# 引用ネットワークの成長性も加味したスコアリング



# 引用ネットワークの成長性も加味したスコアリング



# 引用ネットワークの成長性も加味したスコアリング



# 業種ごとに結果が異なる

## 結果 業種分類：電気機器

スコア	特許の内容
PageRank	レーザープリンタ、DVD、DVDレコーダ
提案手法	広告配信システム、ネット投票システム、情報共有支援システム

- PageRankスコアでは、プリンタ、DVDなど、昔からある重要な製品に関する特許を検出
  - 昔からある&重要な特許＝後発特許から多く引用されている
- 提案手法のスコアでは、情報系など、今後成長が期待される分野の特許を検出
  - 成長性可能性予測が特許の価値評価に寄与



- ✓ 提案手法の方が正確に特許価値を評価
- ✓ 同様の知財戦略を取る他分野でも提案手法のスコアが有効

## 結果 業種分類：医薬品

スコア	特許の内容
PageRank	企業の主力製品となる薬
提案手法	(PageRankほどの特徴なし)

- 業界の性質上、医薬品系特許は物質そのものの特許が多い
  - 類似・代替技術の出願は拒絶されやすい
  - 参入障壁が大きく、成長分野でも技術コミュニティが成長しない
- 発売から時間がたっても、製品の価値が下がりにくい
  - 1つ新薬を出せば、その後長期間にわたって利用される
  - その新薬の関連特許は技術分野の成長性に関わらず価値を維持する



- ✓ PageRankの方が正確に特許価値を評価
- ✓ 手法的に医薬品の分野では提案手法のスコアは有効でない

# 特許・論文



# 異分野横断 技術創造手法

重要  
技術  
特定



機械工学技術

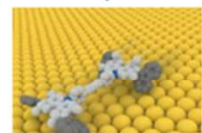
機械工学  
技術



融合



化学  
技術

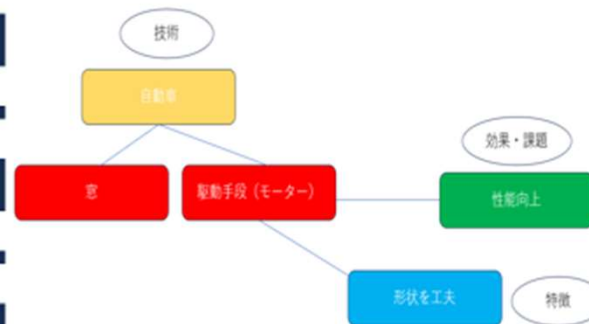


分子マシン

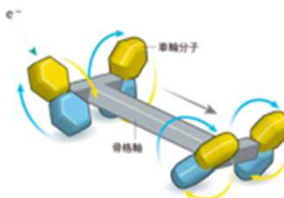
技術融合可能な  
技術群特定

# 創造された技術構成

## 知識グラフ



## 図面



LLMでの技術創造・推論  
(技術構成の具体化)





# 技術の融合可能性の評価

- 技術が融合するかどうかは類似性尺度や課題との共起ネットワークで判断
  - 技術要素としてデータセットに着目した類似性尺度
  - 課題と共起する技術のネットワーククラスタリング

# 技術の融合可能性の評価

- 技術が融合するかどうかは類似性尺度や課題との共起ネットワークで判断
  - 技術要素としてデータセットに着目した類似性尺度
  - 課題と共起する技術のネットワーククラスタリング

IEEE Access  
Multidisciplinary | Rapid Review | Open Access Journal

Received 23 December 2023, accepted 28 February 2024, date of publication 11 March 2024, date of current version 21 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3375750

## RESEARCH ARTICLE

### Metadata-Based Clustering and Selection of Metadata Items for Similar Dataset Discovery and Data Combination Tasks

TAKESHI SAKUMOTO<sup>1</sup>, TERUAKI HAYASHI<sup>2</sup>, HIROKI SAKAJI<sup>3</sup>, AND HIROFUMI NONAKA<sup>4,5</sup>

<sup>1</sup>Department of Engineering, Nagaoka University of Technology, Nagaoka, Niigata 940-2188, Japan

<sup>2</sup>Department of Engineering, The University of Tokyo, Bunkyo, Tokyo 113-8656, Japan

<sup>3</sup>Faculty of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido 060-0814, Japan

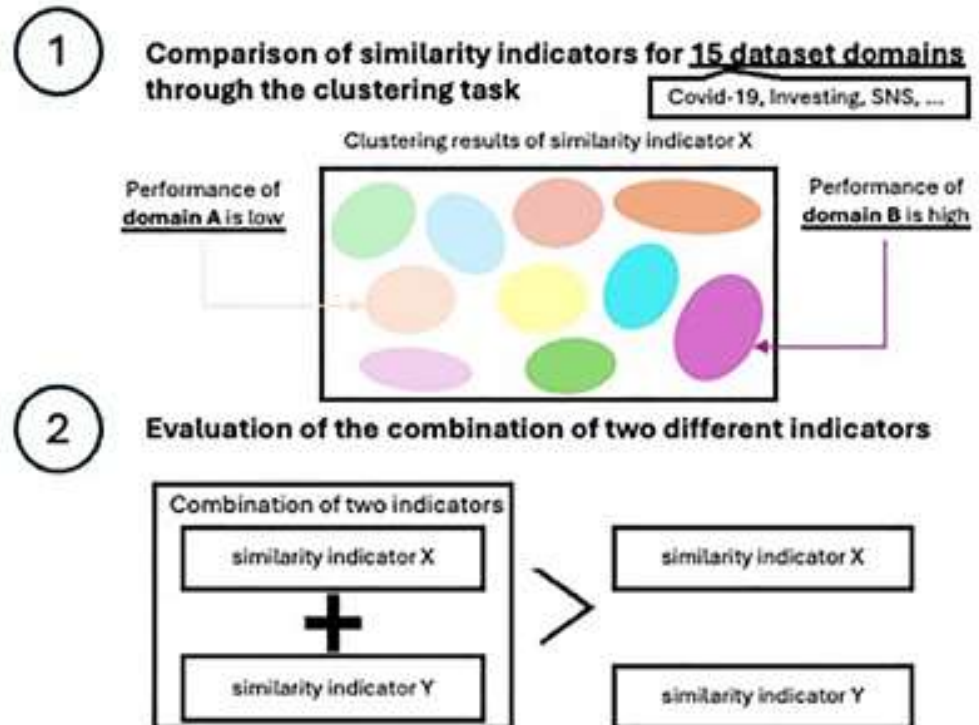
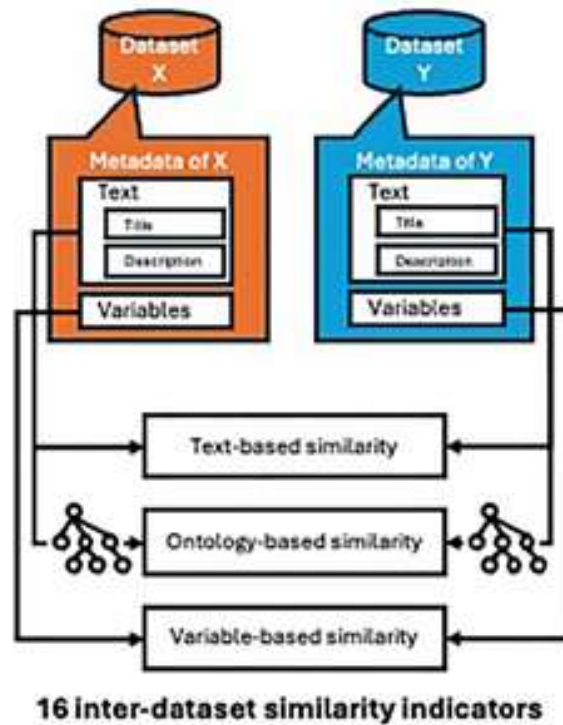
<sup>4</sup>Faculty of Business Administration, Aichi Institute of Technology, Toyota, Aichi 470-0392, Japan

<sup>5</sup>Mayolab Company Ltd., Nagaoka, Niigata 940-2137, Japan

Corresponding author: Takeshi Sakumoto (s183353@stn.nagaokaut.ac.jp)

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant JP20H02384 and Grant JP19K12116.

# データセット融合の際の類似度尺度



# ドメインごとに類似性尺度は異なる

TABLE 5 A Comparison of Clustering Performances for Each Medical Domain

Model	Covid-19	Cancer	Heart Disease
Jaccard(T)	0.288	0.569	0.170
Jaccard(T+D)	0.348	0.471	0.211
Cosine(TF-IDF(T))	<b>0.611</b>	0.582	0.280
Cosine(TF-IDF(T+D))	0.391	0.557	0.330
Cosine(BERT(T))	0.290	0.500	<b>0.406</b>
Cosine(BERT(T+D))	0.173	0.707	0.189
Cosine(Word2Vec(T))	0.205	0.691	0.400
Cosine(Word2Vec(T+D))	0.158	0.626	0.177
DT(Cosine(TF-IDF(T)))	0.417	0.559	0.307
DT(Cosine(TF-IDF(T+D)))	0.504	0.489	0.252
Wu-Palmer(T)	0.220	<b>0.763</b>	0.128
Wu-Palmer(T+D)	0.146	0.738	0.167
Navigational(T)	0.139	0.658	0.169
Navigational(T+D)	0.145	0.494	0.174
Dice(V)	0.238	0.181	0.150
Cosine(TF-IDF(V))	0.292	0.262	0.177

TABLE 6 A Comparison of Clustering Performances for Each Financial Domain

Model	Investing	Currency	Banking
Jaccard(T)	0.414	0.163	0.542
Jaccard(T+D)	0.148	0.168	0.434
Cosine(TF-IDF(T))	0.208	<b>0.703</b>	0.601
Cosine(TF-IDF(T+D))	0.562	0.169	<b>0.646</b>
Cosine(BERT(T))	0.373	0.252	0.331
Cosine(BERT(T+D))	0.208	0.206	0.281
Cosine(Word2Vec(T))	0.538	0.470	0.464
Cosine(Word2Vec(T+D))	0.265	0.212	0.333
DT(Cosine(TF-IDF(T)))	0.500	0.592	0.318
DT(Cosine(TF-IDF(T+D)))	<b>0.578</b>	0.201	0.565
Wu-Palmer(T)	0.370	0.232	0.527
Wu-Palmer(T+D)	0.465	0.210	0.517
Navigational(T)	0.484	0.202	0.390
Navigational(T+D)	0.497	0.180	0.451
Dice(Variables)	0.429	0.311	0.308
Cosine(TF-IDF(Variables))	0.419	0.273	0.211

# 技術の融合可能性の評価

- 技術が融合するかどうかは類似性尺度や課題との共起ネットワークで判断
  - 技術要素としてデータセットに着目した類似性尺度
  - 課題と共起する技術のネットワーククラスタリング

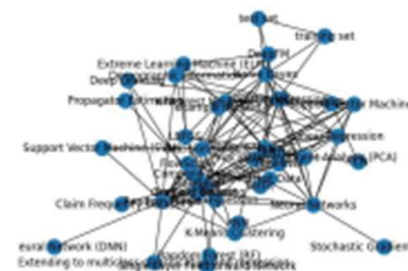
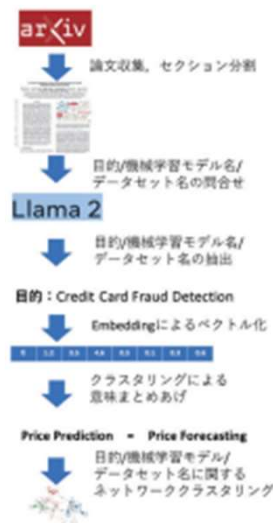
The 38th Annual Conference of the Japanese Society for Artificial Intelligence, 2024

LLM およびネットワーク解析を利用した学術論文からの研究目的・機械学習モデル名・データセット名の抽出と相互の関連性分析  
Extraction of research objectives, machine learning model names, and dataset names from academic papers using large-scale language models and graph network analysis, and a method for analyzing their inter-relationships

西尾 紗也香 \*1    野中 尋史 \*1    早矢仕 晃章 \*2    坂地 泰紀 \*3    作本 猛 \*4  
Sayaka Nishio    Hirofumi Nonaka    Teruaki Hayashi    Hiroki Sakaji    Takeshi Sakumoto

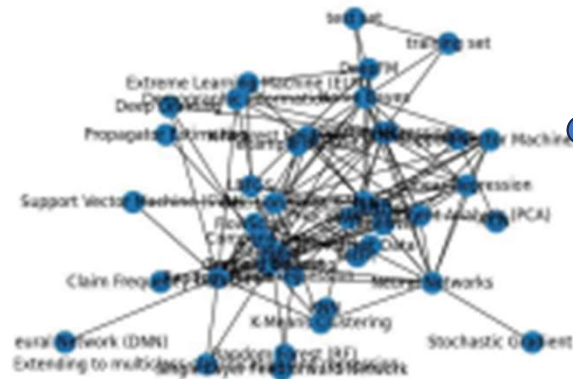
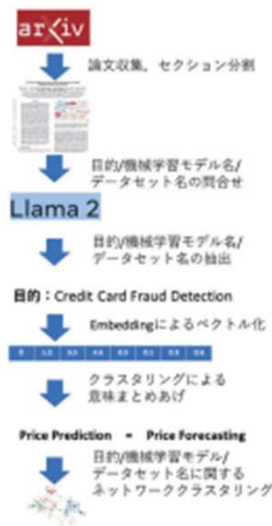
\*1 愛知工業大学    \*2 東京大学    \*3 北海道大学    \*4 長岡技術科学大学  
Aichi Institute of Technology    University of Tokyo    Hokkaido University    Nagaoka University of Technology

In this study, we propose a methodology extracting tasks, machine learning methods, and dataset names from scientific papers and analyzing the relationships between these information.



# 技術の融合可能性の評価

- 技術が融合するかどうかは類似性尺度や課題との共起ネットワークで判断
  - 技術要素としてデータセットに着目した類似性尺度
  - 課題と共起する技術のネットワーククラスタリング



計量ファイナンス分野では予測モデル開発のためにオルタナティブデータの融合が行われようとしている！？

# 特許・論文



## 異分野横断 技術創造手法

重要  
技術  
特定



機械工学技術

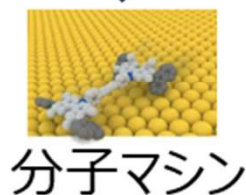
機械工学  
技術



融合



化学  
技術

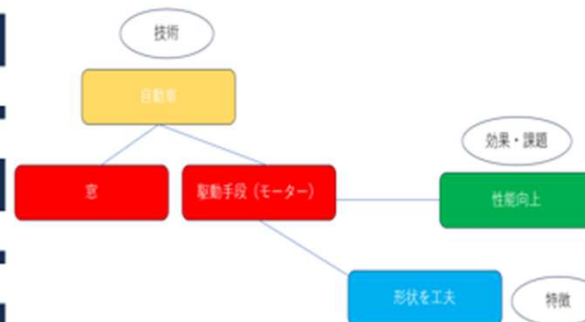


分子マシン

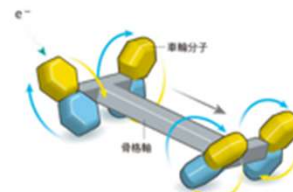
技術融合可能な  
技術群特定

## 創造された技術構成

### 知識グラフ



### 図面



LLMでの技術創造・推論  
(技術構成の具体化)



# 特許・論文



## 異分野横断 技術創造手法

重要  
技術  
特定



機械工学技術

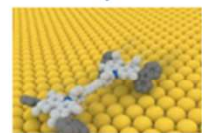
機械工学  
技術



融合



化学  
技術

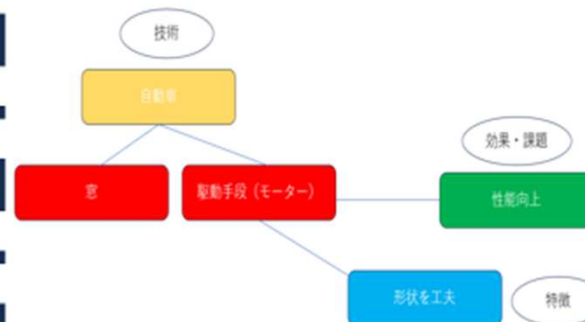


分子マシン

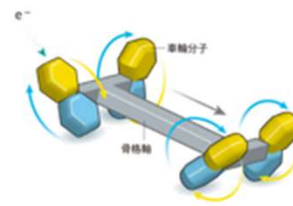
技術融合可能な  
技術群特定

## 創造された技術構成

### 知識グラフ



### 図面



LLMでの技術創造・推論  
(技術構成の具体化)

大規模言語モデルの発展のためには  
良質のデータセット整備が重要

## 科学知識発見を目的とした特許のアノテーション

日浦 隆博<sup>1,2</sup>, 吉田 奈央<sup>3</sup>, 松井 陽子<sup>2</sup>, 河野 誠也<sup>2,1</sup>, 野中 尋史<sup>4</sup>, 吉野 幸一郎<sup>2,1</sup>  
<sup>1</sup> 奈良先端科学技術大学院大学, <sup>2</sup> 理化学研究所 GRP, <sup>3</sup> ANNOTAN, <sup>4</sup> 愛知工業大学  
hiura.takahiro.hu6@is.naist.jp  
{yoko.matsui, seiya.kawano, koichiro.yoshino}@riken.jp  
annotan.tsukuba@gmail.com, hnonaka@aitech.ac.jp

- 特許中の知識を知識グラフとして整理：知識推論へ



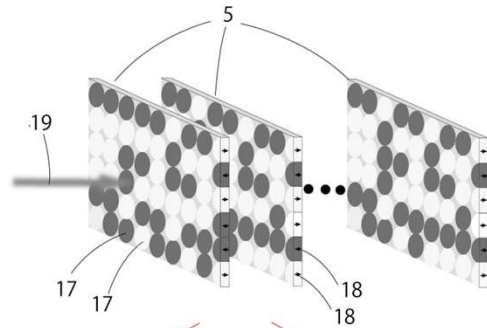
## 特許を対象とする画像言語モデル開発のための データセットの構築

Dataset Development of Vision-Language Model for Patent Data

安藤 七哉<sup>\*1</sup> 溝口 月斗<sup>\*1</sup> 石川 治樹<sup>\*1</sup> 伊豫田 彬<sup>\*1</sup> 河野 誠也<sup>\*2</sup> 吉野 幸一郎<sup>\*2</sup>  
Kazuya Ando Tsukito Mizoguchi Haruki Ishikawa Akira Iyoda Seiya Kawano Koichiro Yoshino

野中 尋史<sup>\*1</sup>  
Hirofumi Nonaka

<sup>\*1</sup>愛知工業大学 Aichi Institute of Technology <sup>\*2</sup>理化学研究所 GRP RIKEN GRP



【課題を解決するための手段】光入力信号が入力する磁性層を備え、該磁性層から光出力信号が出力される磁気光学演算素子であって、光入力信号が直線偏光、円偏光もしくは楕円偏光のいずれかであり、磁性層は磁気光学効果により光出力信号の偏光状態を変える磁気光学材料を含み、磁性層が、光入力信号を推定可能な磁気光学回折ニューラルネットワークを形成する。

図9は磁場印加下におけるレーザー照射による光磁気書き込みの概要図である。

### ○データセットの概要

- ・ 公開特許公報410103件 ( 補正公報除く,2022年-2023年 )
- ・ ケース1(解決手段)とケース2(図のキャプション)を作成
  - 解決手段は特許中のタグ情報から取得
  - 図のキャプションは正規表現を利用

### ○データセット全体の統計情報

- ・ 特許公報当たりの総図面数：12890111件
- ・ 特許公報当たりの平均図面数：31.4枚

### ○ケース1データセット固有の統計情報

- ・ 「課題を解決するための手段」の平均文字数：1506.2文字
- ・ 要約書「解決手段」の平均文字数：236.5文字

### ○ケース2データセット固有の統計情報

- ・ 図のキャプション中の平均文字数：337.5文字

大規模言語モデルの発展のためには  
良質のデータセット整備が重要

# 特許オープンデータ

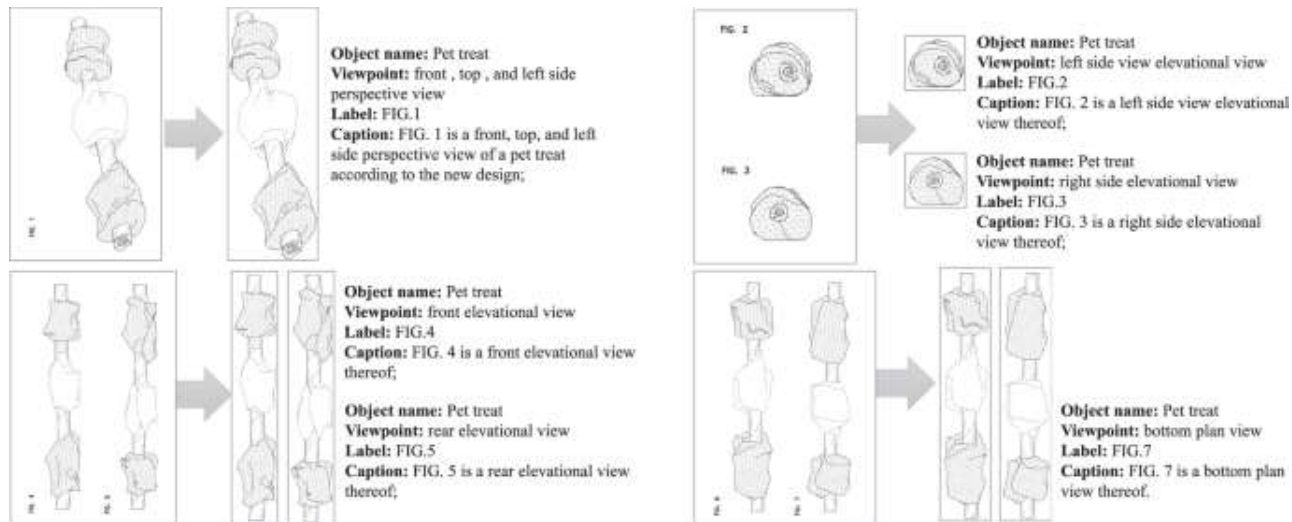
データセット	概要	図面タスクの有無	図面タスクの概要
CLEF-IP 2010年代	英語の特許対象. 検索, 情報抽出など	○	フローチャートとUML の対応
NTCIR PATMN 2000年代	日本 (とUSPTO) の特許対象. 検索, 情報抽出など	×	
TREC-CHEM 2010年代	化学特許対象	○	化学式とテキスト
DeepPatent2 2020年代	米国意匠特許30 万件	○	図と図のキャプション
HUPD	USPTOのテキスト 情報の構造化デー タ	×	

# 特許オープンデータ（図面のみ）

データセット	概要	図面タスクの有無	図面タスクの概要
CLEF-IP 2010年代	英語の特許対象. 検索, 情報抽出など	○	フローチャートとUMLの対応
DeepPatent2 2020年代	米国意匠特許30万件	○	図と図のキャプション
Ours	日本特許40万件	○	図と図に対応する各種 説明文※

# DeepPatent2

<https://www.nature.com/articles/s41597-023-02653-7>



図面とキャプション対応



# 特許関連研究のデータセットの整備について

- 2010年代をピークに減少傾向
- 特許実務者と情報科学分野の研究者が交流し、共同で良質なデータセットを整備する必要がある
  - 特許実務者と情報科学分野の研究者で断絶がある？
  - 官・民・学で連携の必要性

# 知財に関する産学連携

- 牛久先生（ムーンショットプロジェクトPM）、吉野先生、河野先生ら立ち上げた論文・特許情報処理に関する企業「NexaScience」
  - <https://www.nexascience.com/>
- 野崎篤志氏が運営するEパテントチャンネル出演（本日の研究の詳細）
  - <https://www.youtube.com/watch?v=P4iSXebFW7Q>
- 野中が立ち上げた特許解析・モノづくりデータ解析を行う企業「フリーヒルズラボ」
  - [hnonaka@aitech.ac.jp](mailto:hnonaka@aitech.ac.jp) まで（当該企業のことのみならずお気軽にお問合せください）

# 参考文献

- [1] Marusaki, K., Nakai, K., Kataoka, S., Kawano, S., Hentona, A., Sakumoto, T., ... & Nonaka, H. (2024). A study on patent term prediction by survival time analysis using neural hazard model. *Technological Forecasting and Social Change*, 203, 123390.
- [2] Sakumoto, T., Hayashi, T., Sakaji, H., & Nonaka, H. (2024). Metadata-Based Clustering and Selection of Metadata Items for Similar Dataset Discovery and Data Combination Tasks. *IEEE Access*.
- [3] Yamamoto, Y., Hentona, A., Marusaki, K., Watabe, K., Kawano, S., Goto, T., ... & Nonaka, H. (2021, December). Development of the patent values evaluation method considering growth of technical community. In *2021 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1-6). IEEE.
- [4] 日浦隆博, 日浦隆博, 吉田奈央, 松井陽子, 河野誠也, 河野誠也, 野中尋史, 吉野幸一郎, 吉野幸一郎. (2024). 科学知識発見を目的とした特許のアノテーション. *言語処理学会年次大会発表論文集(Web) 30th*.
- [5] 安藤一哉, 溝口月斗, 石川治樹, 伊豫田彬, 河野誠也, 吉野幸一郎, 野中尋史. (2024). 特許を対象とする画像言語モデル開発のためのデータセットの構築. In *人工知能学会全国大会論文集 第38回 (2024)* (pp. 3Xin2103-3Xin2103). 一般社団法人 人工知能学会.
- [6] 西尾紗也香, 野中尋史, 早矢仕晃章, 坂地泰紀, & 作本猛. (2024). LLM およびネットワーク解析を利用した学術論文からの研究目的・機械学習モデル名・データセット名の抽出と相互の関連性分析. In *人工知能学会全国大会論文集 第38回 (2024)* (pp. 3Xin2110-3Xin2110). 一般社団法人 人工知能学会.

リサーチマップのURL

※<https://researchmap.jp/7000022337>

ご清聴ありがとうございました

# 生成AI等を用いた特許文書品質向上のための取り組み

IRD国際特許事務所 所長・弁理士, 博士(情報学)  
大阪大学 招聘教授/京都大学, 関西学院大学 非常勤講師  
谷川 英和  
2024/11/7

# 講師紹介

谷川英和 IRD国際特許事務所 所長 弁理士

## 講師経歴

- 2018年4月～ 関西学院大学非常勤講師(現在に至る)
- 2011年4月～ 大阪大学非常勤講師(現:招聘教授)
- 2009年4月～ 東京工業大学客員教授(～2016年3月)
- 2007年4月～ 京都大学 非常勤講師(現在に至る)
- 2007年3月 京都大学から博士号(情報学)を授与
- 2002年1月 IRD国際特許事務所を開設
- 1999年 弁理士試験合格(登録番号:第11574)
- 1996-2001年 Panasonicの知的財産権部において、デジタルテレビ、ビジネスモデル特許、インターネット関連特許等の多数の権利取得、ライセンス関連業務に従事する。
- 1986～1996年 Panasonicの情報通信研究所でデータベース、ソフトウェア工学の研究開発を行う。自身の研究開発活動において、約180件の特許出願を行っている。
- 1986年 神戸大学工学部システム工学科卒業

## 資格

- 2007年 博士号(情報学)
- 2004年 特定侵害訴訟代理試験合格
- 1999年 弁理士試験合格
- 1994年 情報処理特種技術者試験合格
- 1989年 情報処理第1種技術者試験合格

# 目次

## 1. 生成A Iによる定性特許文書評価

- (1) 特許文書品質向上のためのツール
- (2) 生成A Iを用いた特許文書品質向上のための取り組み
- (3) 産業日本語研究会 特許文書分科会の取り組み

## 2. 機械学習等による定量的特許文書評価

- (1) 機械学習を用いた定量的特許文書評価 (PatentValueAnalyst-SVM)
- (2) 演算式による定量的特許文書評価 (PatentValueAnalyst-CF)

## 3. 特許文書の品質チェック

- (1) 特許文書の品質チェックツール (PatentQC)
- (2) 生成AIを用いた特許文書の品質チェック

## 4. まとめ

# 1. 生成AIによる定性特許文書評価

## (1) 特許文書品質向上のためのツール

生成AIが「業務推進系」にてサポートできるフェーズの検証

今回検証したフェーズ

2-1

2-2

2-3

フェーズ 種類	戦略立案 (企画)	発明構築	特許調査	特許明細書 作成	出願処理	中間処理	登録手続	権利維持	権利無効化	権利行使
業務推進系	特許検索 ツール	発想支援 ツール	特許検索 ツール	明細書 作成支援 システム	出願支援 ツール				特許検索 ツール	
	特許マップ 作成支援 ツール	発明支援 ツール	特許マップ 作成支援 ツール	機械翻訳 システム					文献検索 ツール	
生成AI										
管理系				案件管理 ツール	特許管理ツール			包袋管理ツール		
分析評価系	特許分析 ツール			特許明細書 分析システム						特許分析 ツール
										特許価値 評価 システム



## (2) 生成AIを用いた特許文書品質向上のための取り組み - 生成AIを試用したフェーズ

フェーズ	発明着想	発明構築	特許調査	発明確定	詳細整理	クレーム製造	明細書設計	明細書製造	出願処理	
活動内容	<ul style="list-style-type: none"> <li>課題抽出</li> <li>実施例設計</li> </ul>	<ul style="list-style-type: none"> <li>(R) 本質抽出</li> <li>本質展開</li> </ul>	<ul style="list-style-type: none"> <li>(R) 調査設計</li> <li>(R) 調査</li> </ul>	<ul style="list-style-type: none"> <li>発明確定</li> </ul>	<ul style="list-style-type: none"> <li>(R) 課題等整理</li> </ul>	<ul style="list-style-type: none"> <li>(R) クレーム作成</li> </ul>	<ul style="list-style-type: none"> <li>(R) 明細書構築</li> </ul>	<ul style="list-style-type: none"> <li>(R) 実施例詳細設計</li> <li>(R) 実施例等作成</li> </ul>	<ul style="list-style-type: none"> <li>(R) 願書作成</li> <li>願書事務手続き</li> </ul>	
成果物	発明メモ	発明リスト①	<div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;">調査設計書</div> <div style="border: 1px solid black; padding: 5px;">調査結果書</div>	発明リスト②	出願申請書	クレーム	明細書設計書	図面	明細書	願書など
必要スキル	<ul style="list-style-type: none"> <li>問題提起力</li> <li>技術知識</li> <li>設計力</li> </ul>	<ul style="list-style-type: none"> <li>本質抽出力 (上位概念化力 技術知識)</li> <li>技術知識</li> <li>知識適応力 (モデル適応力)</li> <li>法律知識</li> </ul>	<ul style="list-style-type: none"> <li>本質抽出力</li> <li>技術理解力</li> <li>システム利用能力</li> </ul>	<ul style="list-style-type: none"> <li>本質抽出力</li> <li>技術知識</li> <li>知識適応力</li> <li>法律知識</li> </ul>	<ul style="list-style-type: none"> <li>技術整理力</li> </ul>	<ul style="list-style-type: none"> <li>文章作成能力</li> <li>技術知識</li> <li>設計能力</li> <li>法律知識</li> </ul>	<ul style="list-style-type: none"> <li>技術理解力</li> <li>設計能力</li> <li>法律知識</li> </ul>	<ul style="list-style-type: none"> <li>技術知識</li> <li>設計能力</li> </ul>	<ul style="list-style-type: none"> <li>文書作成能力</li> <li>技術理解力</li> <li>法律知識</li> </ul>	<ul style="list-style-type: none"> <li>法律知識</li> </ul>

・発明起案

・下位概念抽出  
・上位概念抽出  
・発明展開

・特許分類コード決定  
・検索式提案  
・検索

・クレーム作成

・明細書作成(チェック)

・図面作成

(R) レビュー

## (2-1) 発明着想／クレーム製造フェーズ

**[プロンプト]** 犬の散歩のあとの足ふきが面倒なんだけど、何かいい発明ない？

犬の散歩後の足ふきを楽にするための便利なアイデアとして、「自動ペット足洗いマシン」を考えてみます。これは、犬が散歩から帰ったときに使える、簡単かつ効率的な足洗いソリューションです。

自動ペット足洗いマシンの概念: . . . .

**[プロンプト]** その発明について、特許請求の範囲を請求項5つ以内で書いてみて。

特許請求の範囲を考慮して、以下のように請求項を記述することができます。これは、自動ペット足洗いマシンに関する基本的な特徴をカバーするための一例です。

**請求項1:** ペットの足を洗浄するための装置であり、足を挿入するための開口部と、内部に配置された水噴射ノズルおよびブラシを含み、該ブラシは柔らかい素材でできており、ペットの足に優しく作用することを特徴とする装置。

**請求項2:** 請求項1に記載の装置であり、さらにペットの足の挿入を感知するセンサーを含み、該センサーは足の挿入を検知すると自動的に洗浄サイクルを開始することを特徴とする装置。

**請求項3:** 請求項1または請求項2に記載の装置であり、温水および泡を生成する機能を備え、ペットの足を温水と泡で洗浄することができることを特徴とする装置。

**請求項4:** いずれかの請求項に記載の装置であり、足の洗浄後に乾燥サイクルを実行するための乾燥機能を含み、該乾燥機能は温風または自然風の乾燥を提供することを特徴とする装置。

**請求項5:** いずれかの請求項に記載の装置であり、異なるサイズのペットに適應するために開口部の高さまたはサイズを調整する機能を含むことを特徴とする装置。

## (2-2) 発明構築フェーズ [発明展開例]

[プロンプト]以下の発明の下位概念を考えて！ 発明「人の存在を検知し、広告の出力を変更する電気機器」

「人の存在を検知し、広告の出力を変更する電気機器」という発明の下位概念には、より特化したアプリケーションや技術が含まれます。ここでは、それらを実際に適用するいくつかの発明を挙げてみましょう。

**顔認識を用いたデジタルサイネージ**: 通行人の顔の特徴を認識し、年齢や性別に応じたカスタマイズされた広告を表示する。

**動きセンサー搭載ビルボード**: 人の動きを検知して広告を表示し、視線追跡技術を用いて注視点に基づいた広告内容を変更する。

**インタラクティブな店頭ディスプレイ**: 店舗の入口に設置され、近づく顧客に合わせて商品やプロモーション情報を動的に変更する。

**感情認識広告パネル**: 観察された人々の表情を分析し、感情に応じて異なる広告を表示する。

**エンゲージメント測定ディスプレイ**: 観察された人々の関心度合いを測定し、興味を示した内容に関連する広告をさらに表示する。

**パーソナライズされたバーチャルショーケース**: 特定の顧客が商品を見ているときに、その顧客の購買履歴や好みに基づいて商品情報や関連する広告を表示する。

これらの発明は、基本的な概念をさらに具体化し、特定のシナリオや消費者のニーズに合わせて適用される技術です。

## (2-3) 特許調査フェーズ [特許検索式の提案]

[プロンプト]あなたは、優秀な特許のサーチャーです。以下の内容の特許、および以下の内容の特許に類似する特許を検索するための検索式を作成してください。 [発明] 特開2002-140509 の請求項1

特許検索を行う際には、対象となる発明の重要な特徴を考慮して検索キーワードや分類コードを選定する必要があります。ご指定の発明に関して、特許検索式を作成するには、以下のステップに従います。

### 発明の主要特徴の特定

- ・動作が実行できる電気機器
- ・広告情報を蓄積する広告蓄積部
- ・広告情報を取得する広告取得部
- ・広告情報を表示する広告表示部
- ・動作を検知する動作検知部
- ・広告情報の取得と表示を指示する広告変更表示部

### 検索キーワードの選定

- ・広告表示
- ・動作検知
- ・電気機器
- ・広告蓄積
- ・インタラクティブ広告

### 検索式の作成

以下の検索式は、特許データベースにおけるキーワード検索の一例です。使用する特許データベースのインターフェースや機能に応じて調整が必要です。

("電気機器" AND "広告表示" AND "動作検知") OR ("広告蓄積" AND "インタラクティブ広告") OR ("広告取得部" AND "広告変更表示部")

### 分類コードの利用

特許データベースによっては、国際特許分類(IPC)や他の分類システムを利用して検索を行うことができます。この発明に関連する可能性のある分類コードは以下の通りです。

- ・G06F(計算; 計算の推論; 計算装置; データ処理システムまたは方法)
- ・G09F(広告、情報または指示のための装置)
- ・H04N(画像通信, 例. テレビ)

### 注意点

- ・特許検索は複雑であり、検索式は特定のデータベースの機能とユーザーのニーズに応じて調整する必要があります。
  - ・高度な検索では、構文やブール演算子(AND, OR, NOT)を適切に使用することが重要です。
  - ・特許データベースによっては、検索式の構築に更に高度なオプションが用意されている場合があります。
- 最適な検索結果を得るためには、特許データベースの機能を十分に理解し、必要に応じて検索式を調整することが重要です。また、特許専門家によるアドバイスやサポートを求めることも有効です。

## (2-3) 特許調査フェーズ [特許分類コードの提案]

**[プロンプト]**あなたは、優秀な特許のサーチャーです。以下の発明が該当する国際特許分類を教えてください。[発明] 特開2002-140509 の請求項1

この発明は、動作検知に基づいて広告を取得し表示する電気機器に関連するものです。国際特許分類 (IPC) において、この発明が分類される可能性のあるカテゴリーは以下の通りです：

**G06F 3/00** - 計算機システムにおける入力装置または結合装置；これらのための装置。この分類は、動作検知部がユーザーの動作を入力として受け取る装置に該当する可能性があります。

**G09F 19/00** - 広告、情報または指示のための装置や装置の部分。この分類は、広告表示部が含まれるため、発明の広告表示機能に関連しています。

**H04N 21/00** - インタラクティブテレビシステム。この分類は、テレビや類似のディスプレイシステムでのインタラクティブな広告表示機能に関連する可能性があります。

これらの分類は、発明の特定の特徴に基づいて選択されていますが、国際特許分類は非常に広範で詳細なものであるため、この発明の全体的な性質に応じて、他の関連する分類が存在する可能性もあります。特許出願の際には、特許専門家が発明の具体的な特徴に基づいて最も適切な分類を選定します。

<正解>

特開2002-140509のIPC G06F 17/60 G06F 9/00

## (2-3) 特許調査フェーズ [先行特許調査]

[プロンプト] あなたは、優秀な特許のサーチャーです。以下の発明に類似する先行特許を、日本語で教えてください。

[発明] 特開2002-140509 の請求項1

お探しの電気機器に関する先行特許について、具体的な情報を見つけることができませんでしたが、特許検索を効率化するためのツールやサービスがあります。

例えば、Tokkyo.Aiは電気機器業界向けのAI知財生成ツールを提供しており、簡単なアイデアや発明の概要から特許文案を生成でき、AIを用いた類似特許の検索や分析が可能です。これにより市場での競合を避けたり特許を差別化したりすることができるかもしれません。さらに、複数国の特許法に適合した特許申請戦略を策定するのにも利用できます。詳細な先行技術調査や特許検索を行いたい場合は、包括的な特許データベースへのアクセスと専門的な調査能力を持つ特許弁理士や特許検索専門家に相談することをお勧めします。



## (2-4) 明細書製造フェーズ

[明細書を修正する (特許ライティングマニュアルへの適用) ]

基本プロンプト: 以下の特許明細書の日本語を、更に明確にしてください。  
修正の理由も付けて。

改良プロンプト: 以下の特許明細書の日本語を、文字数をあまり増やさずに、  
更に明確にしてください。修正の理由も付けて。

### 【6事例】

1. 長文
2. 複数の主語、述語
3. 目的語無し
4. 主語と述部が離間
5. 修飾語句と被修飾語句とが離間
6. 主語と述語が非対応



# (付録1)

## ひと目で分かる特許ライティングマニュアル

文 レベル	<b>1</b> 短文にする ～短くシンプルな文にする～	一文を短くして複雑な係り受けをなくすことで、人間の理解が容易になり、機械翻訳時の誤訳を減少できる。ただし、短文化する際は、意図しない内容とならないように注意する。
	<b>2</b> 省略しない ～隠れている要素がないか 注意する～	主語や目的語などの省略された要素を明示して、文意を明瞭にする。
節・句 レベル	<b>3</b> 理解しやすい 構成にする ～文の構造に注意する～	文の構造を工夫することで、理解しやすく翻訳しやすい文にすることができる。
	<b>4</b> 横並びの要素の 表現を揃える ～対等に並べ意味に注意する～	要素を対等に並べることで文章に構造をもたせ、係り受けを明確にし、理解しやすくする。
	<b>5</b> 読点を工夫する ～係り受けや文の構造を 明らかにする～	適切な位置に読点をつけることで、係り受けや文の構造を明らかにする。
	<b>6</b> 簡潔にする ～シンプルな表現にする～	不要、冗長、難解な表現を避け、元の文の意味が変わらない簡潔な表現を用い、シンプルな文にする。
語 レベル	<b>7</b> 言い換える ～誰にでも伝わる表現にする～	多義的な表現やあいまいな表現は、明確・具体的な表現に改める。日本語特有の表現は、訳しやすい表現に改める。

一般財団法人日本特許情報機構 (Japio) は、2007年から、知的財産、特許翻訳、情報工学等の専門家の皆様のご協力のもとで、「産業日本語」の研究・普及活動を行っております。その活動の中で、「言い換えルール」の抽出を行いました。この言い換えルールを一般にも親しみやすく理解しやすいものとなるよう、8つのカテゴリー、31のルールに汎用化したものを、「特許ライティングマニュアル（初版）」として、2013年6月に発行しました。

## (2-4) 明細書製造フェーズ [明細書を修正する]

### 事例1:長文

修正前:

耐熱ガラスからなる透明な表面側第1層、充填材からなる第2層、複数の光電変換素子が配置された第3層、樹脂フィルムからなる耐候性の裏面側第4層をこの順序で積層する。

【JAPIO修正】:

表面側第1層、第2層、第3層、裏面側第4層をこの順序で積層する。表面側第1層は耐熱ガラスからなり透明である。第2層は充填材からなる。第3には複数の光電変換素子が配置されている。裏面側第4層は樹脂フィルムからなり耐候性である。

【GPT-4修正】:

耐熱性を有するガラス製の透明な表面側第1層、充填材によって形成された第2層、複数の光電変換素子が組み込まれている第3層、そして耐候性を備えた樹脂フィルムで構成される裏面側第4層を、この指定された順序に従って積層する。

コメント:

GPT-4はJAPIOテキストのように5文に区切らず、オリジナルの1文の構成を維持した上で、各層の修飾方法を変更しているため、JAPIO修正よりも明確性を向上させる修正とは評価しにくい。

各層の修飾は、原文にないグレーゾーンの情報による冗長性が追加されている印象。

## (2-4) 明細書製造フェーズ [明細書を修正する]

### 事例2: 複数の主語、述語

修正前:

積層フィルムの層間にアルミの第1層、酸化ケイ素の第2層を設けることで、積層フィルムの水蒸気バリア性を確保できるため、食品、医薬品等の内容物の包装に好適である。

【JAPIO修正】:

積層フィルムの層間にアルミの第1層、酸化ケイ素の第2層を設けることで、積層フィルムの水蒸気バリア性を確保できる。**このため、かかる積層フィルムは、**食品、医薬品等の内容物の包装に好適である。

【GPT-4修正】:

積層フィルムの各層間に、アルミニウムを用いた第1層と酸化ケイ素による第2層を導入することにより、当該積層フィルムの**水蒸気バリア性が強化される**。この結果、食品や医薬品などの内容物を包装する際に、その品質を保持するのに特に適している。

コメント:

GPT-4はJAPIO修正と同様、2文に区切って明確性を改善している。

更に、GPT-4の「**水蒸気バリア性が強化される**」への修正は、文法的に**能動形を受動形に変更し**

**主語なしの文を明確にしており、JAPIO修正を更に改善している。**

但し、原文にないグレーゾーンの情報による冗長性が追加されている印象。

## (2-4) 明細書製造フェーズ [明細書を修正する]

### 事例3: 目的語無し

修正前:

洗浄ノズルは、高圧の洗浄水を噴射して洗浄する。

【JAPIO修正】:

洗浄ノズルは、高圧の洗浄水を噴射して**汚染部**を洗浄する。

【GPT-4修正】:

本実施例における洗浄ノズルは、高圧で噴射される洗浄水を用いて**対象物**を効率的に洗浄する機能を有する。

コメント:

GPT-4はJAPIO修正と同様、**目的語無しの問題**に対し、目的語を追加して適切な修正。

但し、原文にないグレーゾンの情報による冗長性が追加されている。

## (2-4) 明細書製造フェーズ [明細書を修正する]

### 事例4: 主語と述語が離間

修正前:

第1制御部は、温度検出器により検出された油の温度が設定温度より低い場合には、電源ケーブルの途中に介装されたサイリスタ、トライアック等からなる通電制御部に通電信号を出力する。

【JAPIO修正】:

温度検出器により検出された油の温度が設定温度より低い場合には、**第1制御部は**、電源ケーブルの途中に介装されたサイリスタ、トライアック等からなる通電制御部に通電信号を出力する。

【GPT-4修正】:

第1制御部は、温度検出器によって検出された油の温度が設定温度未満である場合、電源ケーブルに組み込まれているサイリスタやトライアックなどの通電制御部分に対し通電を促す信号を出力する。この通電制御部分は、電源ケーブルの途中に設置されている。

コメント:

GPT-4は一番の論点の、**主語と述部が分離された部分を近づける修正がされていない。**

GPT-4の修正は、情報を冗長にする修正でJAPIO案よりも改善無し。

## (2-4) 明細書製造フェーズ [明細書を修正する]

### 事例5: 修飾語句と被修飾語句とが離間

修正前:

ロボットアームは部品と商品搬送時に接触する。

【JAPIO修正】:

ロボットアームは商品搬送時に**部品**と接触する。

【GPT-4修正】:

本実施例におけるロボットアームは、**部品の取り扱いおよび商品の搬送プロセス中に**、これらの**物体に直接接触**して作業を行う。

コメント:

GPT-4はJAPIO修正の、「部品と接触する」とは異なり、「物体に直接接触して」と、記載が異なる修正に

見える。

しかし手前で、“**部品**”および“**商品**”が記載されており、原文の意味からシフトしており、**不適切な修正**。

更に、原文にない情報を追加して冗長性を増しており、JAPIO修正よりも改善無し。

## (2-4) 明細書製造フェーズ [明細書を修正する]

### 事例6: 主語と述語が非対応

修正前:

自走ロボットの停止点は、地点A、地点B、地点Cに止まり、その後出発点に戻る。

【JAPIO修正】:

自走ロボットは、地点A、地点B、地点Cに止まり、その後出発点に戻る。

【GPT-4修正】:

本実施例における自走ロボットは、プログラムされたルートに従い、地点A、地点B、および地点Cで順次停止する。各停止地点での任務を完了した後、ロボットは出発点に自動的に戻る。

コメント:

GPT-4は、原文の主語と述語とのねじれとは、適切に修正している。

但し、今回の修正のスタイルのように、原文にない不必要な情報を追加しており、適切とは言い難い。

## (2-4) 明細書製造フェーズ [明細書を修正する]

### 事例2(改良プロンプト): 複数の主語、述語

【JAPIO修正】:

積層フィルムの層間にアルミの第1層、酸化ケイ素の第2層を設けることで、積層フィルムの水蒸気バリア性を確保できる。**このため、かかる積層フィルムは、食品、医薬品等の内容物の包装に好適である。**

【GPT-4修正】:

積層フィルムの各層間に、アルミニウムを用いた第1層と酸化ケイ素による第2層を導入することにより、当該積層フィルムの**水蒸気バリア性が強化される**。この結果、食品や医薬品などの内容物を包装する際に、その品質を保持するのに特に適している。

【GPT-4改良プロンプト修正】

アルミ製第1層と酸化ケイ素第2層を積層フィルム間に配置し、水蒸気バリア性を向上。これにより**食品や医薬品包装に最適**。

コメント:

改良プロンプトは体言止めとなり、文字数削減が優先されるが書式として不適切。

原文の「食品、医薬品等の内容物の包装に好適である。」の中で、「内容物」が削られ情報がスリムになった。

**クレーム作成時に、 unnecessary要件を削ぎ落とした、独立項の作成に活用できる可能性がある。**



## (2-4) 明細書製造フェーズ [明細書を修正する]

### 事例3(改良プロンプト): 目的語無し

【JAPIO修正】:

洗浄ノズルは、高圧の洗浄水を噴射して汚染部を洗浄する。

【GPT-4修正】:

本実施例における洗浄ノズルは、高圧で噴射される洗浄水を用いて対象物を効率的に洗浄する機能を有する。

【GPT-4改良プロンプト修正】:

洗浄ノズルは、高圧水を噴射し対象を効率的に洗浄。

コメント:

GPT-4改良プロンプトはJAPIO修正とほぼ同様の修正案を作成。

不足する目的語を「対象を」と適切に追加した上で、余計な情報を追加しない修正案。但し、体言止めは、文字数制限の条件から発生しており、別のプロンプトが必要。

## (2-4) 明細書製造フェーズ [明細書を修正する]

### 事例6(改良プロンプト): 主語と述語が非対応

【JAPIO修正】:

自走ロボットは、地点A、地点B、地点Cに止まり、その後出発点に戻る。

【GPT-4修正】:

本実施例における自走ロボットは、プログラムされたルートに従い、地点A、地点B、および地点Cで順次停止する。各停止地点での任務を完了した後、ロボットは出発点に自動的に戻る。

【GPT-4改良プロンプト修正】:

自走ロボットは地点A、B、Cで停止後、出発点へ戻る。

コメント:

GPT-4改良プロンプトはJAPIO修正とほぼ同様の修正案を作成。

主語と述語のねじれを解消し、更には、JAPIO修正を究極にスリム化している。

本事例は体言止めは発生しておらず、完成度の高い修正案になっている。

## (2-4) 明細書製造フェーズ [明細書を修正する]

<プロンプトB>

あなたは優秀な弁理士です。

三重引用符で囲われた明細書の一部の文章を校閲し、修正を行った場合には修正箇所に対応する理由を示してください。また、文章として誤っているが、与えられている情報から修正案が作成できない場合には、誤っている点を指摘してください。

その際には、以下の点を改善することを意識して校閲を行ってください。

- ・ **正確性**

用語または文が文法的に正確である度合い。例えば、誤記が多いほど、主語の抜けが多いほど、用語の揺らぎの数が多いほど、評価は低くなる。

- ・ **簡潔性**

文または文章が簡潔である度合い。例えば、多義的に解釈され得る用語または文が多いほど、評価は低くなる。

- ・ **一義性**

用語または文が一義的に捉えられる度合い。例えば、多義的に解釈され得る用語または文が多いほど、評価は低くなる。

- ・ **技術開示性**

特許請求の範囲に記載された発明が十分に説明されている度合い。なお、サポート要件、実施可能要件を満たす度合いとも言える。例えば、特許請求の範囲を構成する用語の技術的意義、構成と作用と効果が十分に説明されていないほど、実施例や変形例が不十分であるほど、評価は低くなる。

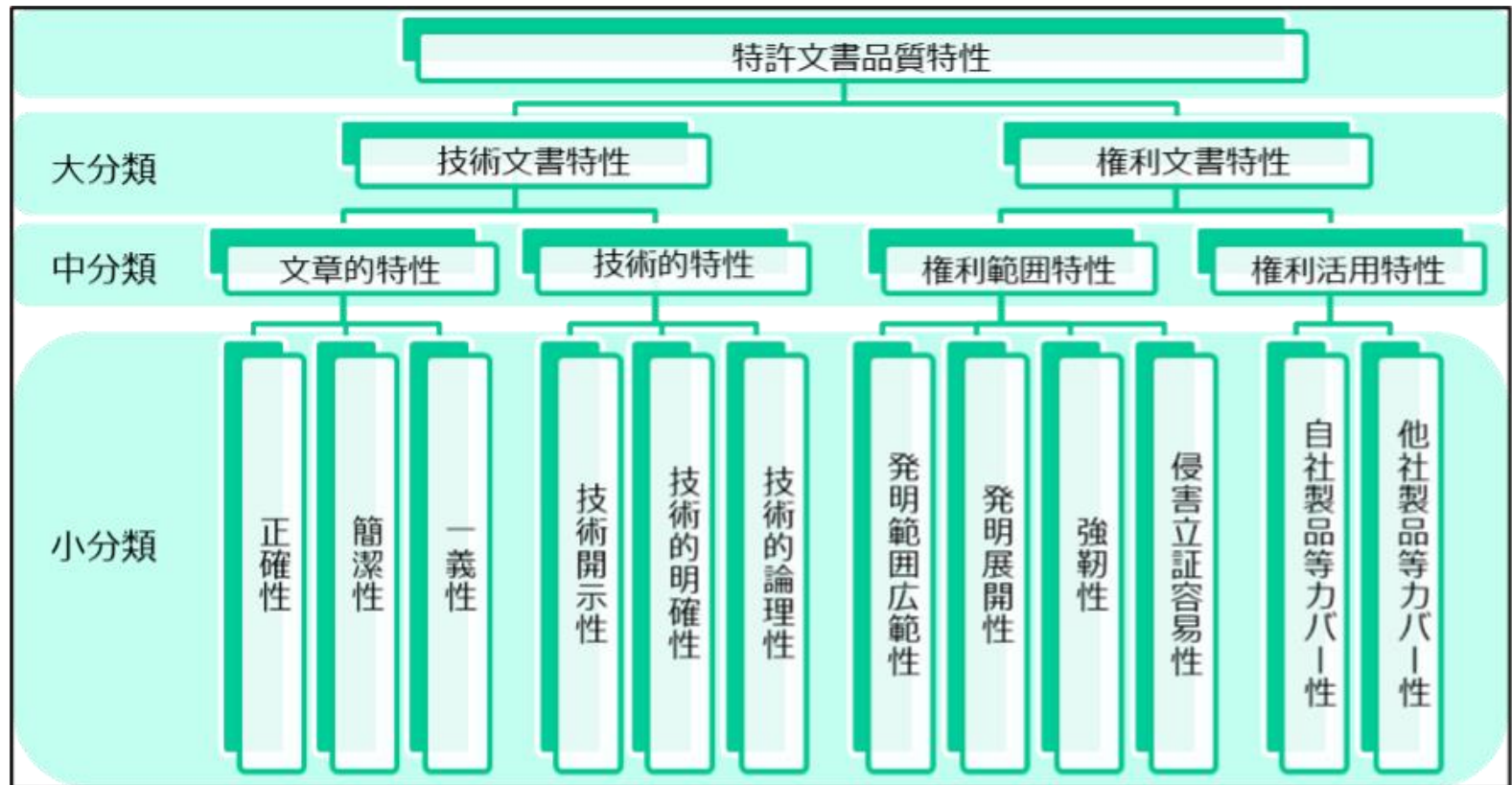
- ・ **技術的明確性**

発明が技術的に分かりやすく明確に説明されている度合い。例えば、構成と作用の効果の記載が不十分であるほど、上位概念の用語または造語に対して、具体例が少ないほど、評価は低くなる。

- ・ **技術的論理性**

発明が技術的に論理的に説明されている度合い。例えば、次の場合は、評価が低くなる。①クレームと明細書のストーリーとの技術的因果関係が不明確である、②従来技術、課題、解決手段、効果のストーリーの筋が通っていない、③実施例の技術説明に論理的な誤りが多い。

# 特許文書品質特性モデル



## 【悪例】

なお、**砒素(P)**は同時にゲートポリシリコン膜104にも注入され、ゲートポリシリコン膜104の導電率を高める。



## 【良例】

なお、**砒素(As)**は同時にゲートポリシリコン膜104にも注入され、ゲートポリシリコン膜104の導電率を高める。

<プロンプトBの回答1>

修正案:

"""

なお、砒素 (As)は同時にゲートポリシリコン膜104にも注入され、ゲートポリシリコン膜104の導電率を高める。

"""

理由:

元の文は文法的に正しいです。ただし、砒素の元素記号が誤っています。正しくは**As**であるため、正確性と一義性を修正しました。

<プロンプトBの回答2>

修正案:

"""

なお、砒素 (As)は同時にゲートポリシリコン膜104にも注入され、ゲートポリシリコン膜104の導電率を高める。

"""

修正理由:

砒素の化学記号は正確には**(As)**であり、**(P)**ではありません。したがって、正確性を確保するために、この部分を修正しました。

(2) 生成AIを用いた特許文書品質向上のための取り組み  
 - 生成AIを試用したフェーズ

フェーズ	発明着想	発明構築	特許調査	発明確定	詳細整理	クレーム製造	明細書設計	明細書製造	出願処理	
活動内容	<ul style="list-style-type: none"> <li>課題抽出</li> <li>実施例設計</li> </ul>	<ul style="list-style-type: none"> <li>(R) 本質抽出</li> <li>本質展開</li> </ul>	<ul style="list-style-type: none"> <li>(R) 調査設計</li> <li>(R) 調査</li> </ul>	<ul style="list-style-type: none"> <li>発明確定</li> </ul>	<ul style="list-style-type: none"> <li>(R) 課題等整理</li> </ul>	<ul style="list-style-type: none"> <li>(R) クレーム作成</li> </ul>	<ul style="list-style-type: none"> <li>(R) 明細書ストーリー構築</li> </ul>	<ul style="list-style-type: none"> <li>(R) 実施例詳細設計</li> <li>(R) 実施例等作成</li> </ul>	<ul style="list-style-type: none"> <li>(R) 願書作成</li> <li>願書事務手続き</li> </ul>	
成果物	発明メモ	発明リスト①	<div style="border: 1px solid black; padding: 5px;">調査設計書</div> <div style="border: 1px solid black; padding: 5px;">調査結果書</div>	発明リスト②	出願申請書	クレーム	明細書設計書	図面	明細書	願書など
必要スキル	<ul style="list-style-type: none"> <li>問題提起力</li> <li>技術知識</li> <li>設計力</li> </ul>	<ul style="list-style-type: none"> <li>本質抽出力 (上位概念化力 技術知識)</li> <li>技術知識</li> <li>知識適応力 (モデル適応力)</li> <li>法律知識</li> </ul>	<ul style="list-style-type: none"> <li>本質抽出力</li> <li>技術理解力</li> <li>システム利用能力</li> </ul>	<ul style="list-style-type: none"> <li>本質抽出力</li> <li>技術知識</li> <li>知識適応力</li> <li>法律知識</li> </ul>	<ul style="list-style-type: none"> <li>技術整理力</li> </ul>	<ul style="list-style-type: none"> <li>文章作成能力</li> <li>技術知識</li> <li>設計能力</li> <li>法律知識</li> </ul>	<ul style="list-style-type: none"> <li>技術理解力</li> <li>設計能力</li> <li>法律知識</li> </ul>	<ul style="list-style-type: none"> <li>技術知識</li> <li>設計能力</li> </ul>	<ul style="list-style-type: none"> <li>文書作成能力</li> <li>技術理解力</li> <li>法律知識</li> </ul>	<ul style="list-style-type: none"> <li>法律知識</li> </ul>

・発明起案

・下位概念抽出  
 ・上位概念抽出  
 ・発明展開

・特許分類コード決定  
 ・検索式提案  
 ・検索

・クレーム作成

・明細書作成(チェック)

・図面作成

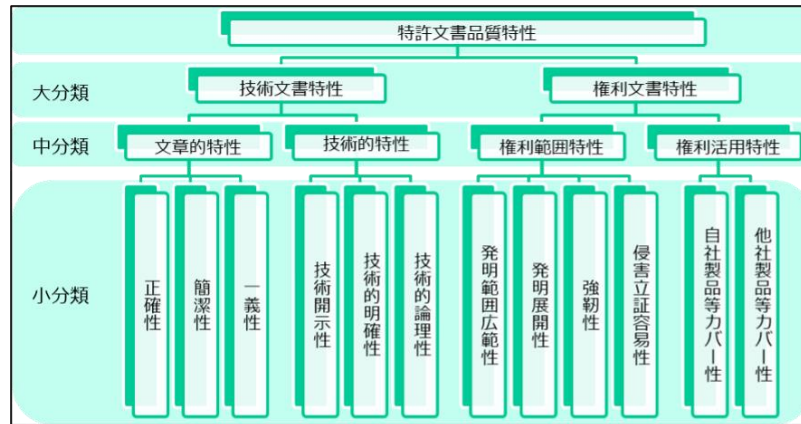
(R) レビュー

(3) 産業日本語研究会 特許文書分科会の取り組み  
(3-1) 背景と活動概要

特許文書分科会の目的：特許文書の品質向上に寄与する

高品質な特許文書とは

特許文書  
品質特性モデル



特 許 ラ イ フ サ イ ク ル

発明着想 発明完成 特許明細書作成 出願 中間処理 権利化 権利行使

教育

方法論  
(成果物・プロセス)

ツール

## (3) 産業日本語研究会 特許文書分科会の取り組み (3-2) 教育 [特許文書品質の教育用テキスト]

# 特許文書品質特性モデルの学習用テキスト(β版)を 産業日本語研究会HP (tech-jpn.jp)にて公開[2023年3月6日]

### 特許文書品質特性モデルの学習用テキスト

#### ◎特許文書品質特性モデルの学習用テキストとは

特許の権利化、特許権の行使等を鑑みれば、特許請求の範囲、明細書等の特許文書の品質が重要となりますが、これまで、特許文書の「品質」についての研究や深い議論がなされた文献が見当たりませんでした。

産業日本語研究会「特許文書分科会」では、特許文書の「品質」を標準化および定量化すべく、平成28年度より特許文書品質特性モデルについて検討を行い、この度、特許文書品質特性モデルの学習用テキストを公開いたしました(本ページ下部のダウンロードボタンよりダウンロード可能です)。

本学習用テキストでは、特許文書品質特性モデルの利用方法や、品質特性ごとの良例・悪例等について学ぶことができるようになっています。

より高品質な特許文書を作成するための自主学習や研修のテキストとして利用いただくなど、ご自由に活用ください。また、当該テキストをより良いものとするため、皆様からのご意見を頂けると幸いです。

#### 更新履歴

- α版公開 2022年11月4日
- β版公開 2023年3月6日(全体構成の見直し、特許文書品質チェックリストの作成、記載内容の見直し等)

#### 学習用テキストの目次

- 第1章 特許文書品質特性の背景
  - 第2章 特許文書品質特性の定義・概要
  - 第3章 特許文書品質特性の利用方法
    - 3. 1 特許文書特性の評価方法
    - 3. 2 定性的な評価方法
    - 3. 3 定量的な評価方法
    - 3. 4 特許文書の品質評価の場面
  - 第4章 分野別の特許文書品質特性の事例(悪例→良例)
    - (4. 1) 機械 (4. 2) 電気 (4. 3) 化学 (4. 4) ソフト
  - 第5章 分野別のサンプル明細書と特許文書品質特性との関係
    - (5. 1) 機械 (5. 2) 電気 (5. 3) 化学 (5. 4) ソフト
  - 第6章 演習課題(悪例を多く含むサンプル明細書に基づく)
    - (6. 1) 機械 (6. 2) 電気 (6. 3) 化学 (6. 4) ソフト
  - 第7章 終わりに
- (付録)
- (A) 特許文書品質チェックリスト(詳細版, 簡易版)
  - (B) サンプル明細書( (1) 機械 (2) 電気 (3) 化学 (4) ソフト) [別ファイル]
  - (C) 特許文書品質特性モデルのしおり [別ファイル]

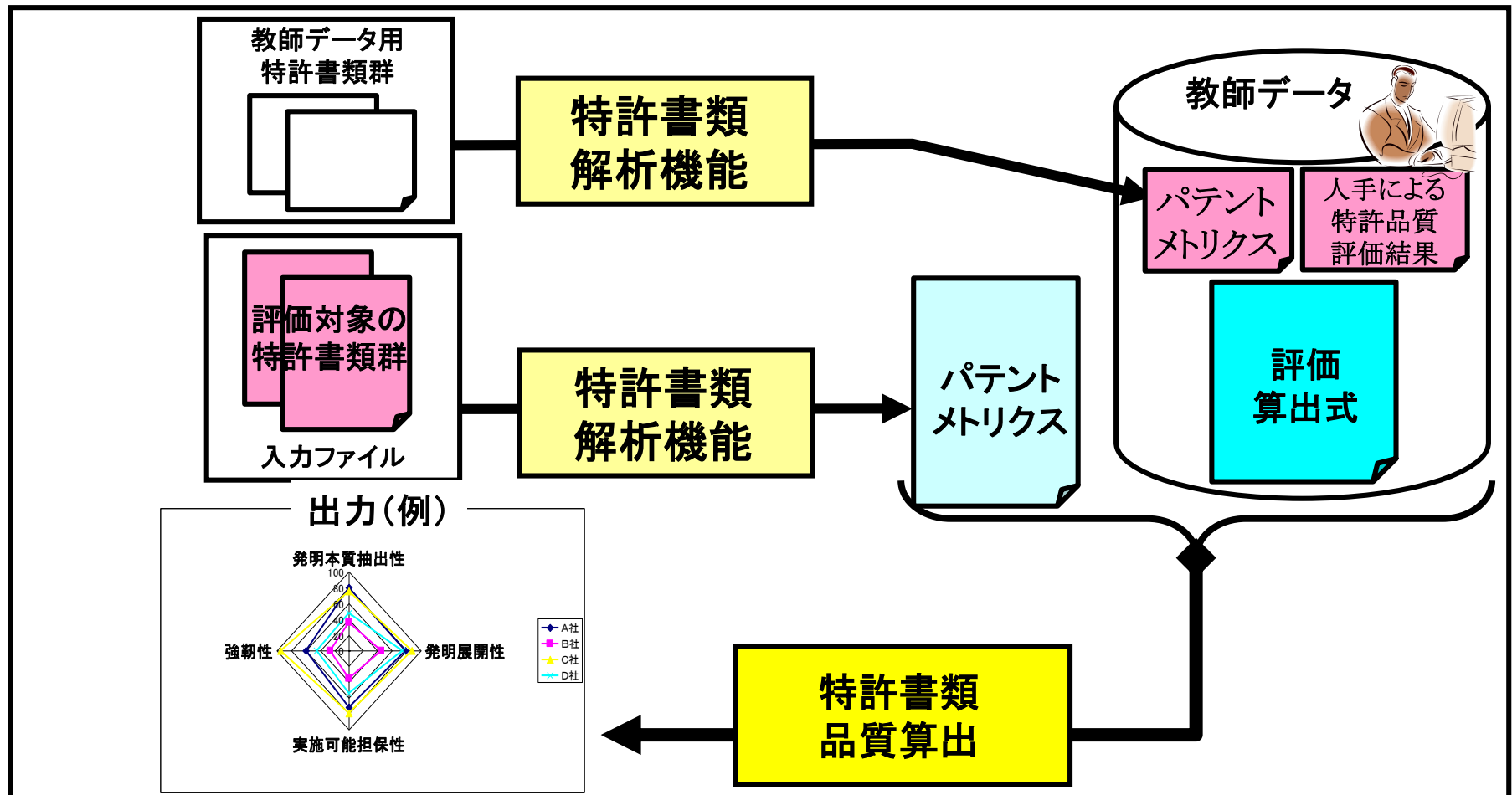
(URL) <https://tech-jpn.jp/tokkyo-hinshitu-gakusyu-text/>



## 2 機械学習等による定量的特許文書評価

- (1) 機械学習を用いた定量的特許文書評価 (PatentValueAnalyst-SVM)
- (2) 演算式による定量的特許文書評価 (PatentValueAnalyst-CF)

品質評価対象の特許書類を言語解析し、取得したパテントメトリクスを用いた品質の自動算出  
[評価式算定方式] 予め決められた評価算出式に代入し、品質を自動算出する  
[機械学習方式] 過去の評価実績の機械学習により、対象特許の品質を自動算出する



# 約60種類のパテントメトリクス

No	要素名	属性名	意味	品質との相関性	説明・備考
1	patent-analyzed	date-of-analyzed	解析日	—	PMLファイルを作成した日
2		duration	存続期間	—	解析日から算出した存続期間
3	claims	number-of-claims	請求項の数		
4		number-of-independent-claims	独立形式請求項の数	↑	
5		number-of-categories	カテゴリーの数	↑	
6		nest-level	ネストレベル	↑	ネストレベルとは、請求項の従属関係を木構造で表した場合の木構造の深さのこと 木の根が複数存在する場合は、最も深いレベル
7		number-of-characters	文字数の合計	→	
8		number-of-claims-of-representation-category	代表カテゴリーに属する請求項数	↑	代表カテゴリーとは、最も請求項数が多いカテゴリーのこと 同数のカテゴリーが2以上存在する場合、同一のカテゴリーの中で小さい番号の請求項のカテゴリーを代表カテゴリーとする
9		number-of-independent-claims-of-representation-category	代表カテゴリーに属する独立形式請求項数	↑	
10		nest-level-of-representation-category	代表カテゴリーに属する請求項のネストレベル	↑	
11		number-of-characters-of-representation-category	代表カテゴリーに属する請求項の文字数の合計	→	
12		number-of-vague-terms	曖昧な用語の数	↓	「除くでない以上 以下 やや はるか 高温 低温 高圧 低圧 にくい やすい 所望により必要により特に 例えば など 好ましくは 適宜」の数
13		number-of-purpose-terms	目的を示す用語の数	↓	「ため できる」の数
14		number-of-not-explained-elements	実施の形態において説明されていない構成要素の数	↓	
15		number-of-not-explained-terms	実施の形態において説明されていない用語の数	↓	
16		number-of-not-fully-explained-elements	実施の形態において十分に説明されていない構成要素の数	↓	
17		number-of-not-fully-explained-terms	実施の形態において十分に説明されていない用語の数	↓	
18		number-of-inappropriate-dependent-claims	不適切な引用形式請求項の数	↓	存在していない請求項を引用している請求項、自身よりも番号の大きい請求項を引用している請求項の数
19	claim-info	claim-title	請求項の名称	—	請求項本文の最後の「情報処理装置」、「製造方法」、「積層型鉄心の製造方法」などの名詞句
20		type1	独立形式請求項か、引用形式請求項か	—	独立形式請求項:「請求項1記載の」などの従属表現がない ⇒independent 引用形式請求項:「請求項1記載の」などの従属表現がある ⇒dependent
21		type2	ジェブソンの形式か、順次列挙形式か、構成要素列挙形式か	—	ジェブソンの形式:冒頭に「～において、」や「～であって、」があるもの ⇒jepson 順序列挙形式:「～して、」や「～し、～し、～し」などの形式になっているもの ⇒processSequence 構成要素列挙形式:「～と、～と、～とから構成される○○」などの形式になっているもの ⇒elementEnumeration ジェブソンの形式かつ順序列挙形式の場合: ⇒jepson.elementEnumeration ジェブソンの形式かつ構成要素列挙形式の場合:
22		parent	親クレーム(従属元の請求項)	—	「請求項1記載の」などの番号
23		category	カテゴリー	—	請求の名称の最後が 「方法」:method 「プログラム」:program 「媒体」:media その他:device
24		number-of-characters	請求項の文節数	→	
25		number-of-phrases	請求項の文節数	→	
26		number-of-characters-of-premise	前提部の文字数	→	
27		number-of-characters-of-feature	特徴部の文字数	→	
28		number-of-phrases-of-premise	前提部の文節数	→	
29		number-of-phrases-of-feature	特徴部の文節数	→	
30		number-of-vague-terms	曖昧な用語の数	↓	「除くでない以上 以下 やや はるか 高温 低温 高圧 低圧 にくい やすい 所望により必要により特に 例えば など 好ましくは 適宜」の数
31		number-of-purpose-terms	目的を示す用語の数	↓	「ため できる」の数
32		illegal-quotation	不適切な引用形式かどうか	—	不適切な引用形式の場合:存在していない請求項を引用している、または自身よりも番号の大きい請求項を引用している ⇒yes 不適切な引用形式でない場合 ⇒no
33	description	number-of-quoted-patents	<背景技術、従来技術>引用されている特許文献数	↑	「[背景技術]」の欄の「[特許文献#]」の数
34		number-of-quoted-non-patents	<背景技術、従来技術>引用されている非特許文献数	↑	「[背景技術]」の欄の「[非特許文献#]」の数
35		number-of-embodiments	実施の形態の数	→	「[背景技術]」の欄の「[実施の形態#]」の数
36		number-of-examples	実施例の数	→	「[背景技術]」の欄の「[実施例#]」の数
37		number-of-embodiment-characters	<発明の実施の形態>「発明の実施の形態」の文字数	→	「[発明を実施するための形態]」の欄の文字数
38		number-of-embodiment-phrases	<発明の実施の形態>「発明の実施の形態」の文節数	→	「[発明を実施するための形態]」の欄の文節数
39		number-of-example-characters	<発明の実施の形態>「実施例」の文字数	→	「[実施例]」の欄の文字数
40		number-of-example-phrases	<発明の実施の形態>「実施例」の文節数	→	「[実施例]」の欄の文節数

### 3. 特許文書の品質チェック

## (1) 特許文書の品質チェックツール (PatentQC)

### サポート要件のチェック(第36条第6項第1号)

構成要素のチェック, 用語のチェック

### 明確性要件のチェック(第36条第6項第2号)

未記載用語のチェック, 不適切な引用関係のチェック  
曖昧語のチェック, 目的語のチェック

### 翻訳容易性・可読性のチェック

⇒明細書の可読性(読みやすさ)に関するチェック

複雑な文のチェック, 主語のない文のチェック,  
閾値以上の文字数の文のチェック

# (1) 明細書チェックツール (PatentQC)

特許出願書類のチェック

入力ファイル  
明細書.docx

出力ファイル  
明細書\_チェック済.docx

サポート要件のチェック

- 構成要素名 1000 文字以下
- 用語 500 文字以下

明確性要件のチェック

- 曖昧語  未記載用語
- 目的語
- 不適切な引用関係

可読性のチェック

- 指定文字数以上の文 150 文字以上
- 指定文字数以上の段 500 文字以上
- 主語のない文
- 複雑な文
  - 絶対的な複雑度 50 以上(1~)
  - 相対的な複雑度 50 以上(1~100)

チェック開始 終了

Version: 1.0.1

## (2) 生成AIを用いた明細書チェック

仮想明細書（タクシー）を題材として、以下の指摘がChatGPT4o（以下「GPT」）に可能であるか否かを検査した。

1. クレームの従属関係
  - (1) マルチ・マルチクレームの検出
  - (2) 従属関係の不備の検出
2. 技術的明確性
  - (1) 「前記・・・」の不備の検出
  - (2) 不明確なクレームの検出
3. 文章的特性（正確性・簡潔性）
  - (1) 主語の無い文の検出／修正
  - (2) 複文の検出／修正
  - (3) 長文の検出／修正

【書類名】特許請求の範囲

【請求項 1】

基地局、顧客端末、および車載装置からなるタクシー捕捉システムにおいて、前記基地局は、複数の空車の車載装置から受信したそれぞれの空車の位置情報を格納する空車位置データベースと、地図を記憶する地図データベースと、前記顧客端末から受信した顧客の位置情報に基づいて、該顧客の周囲に存在する一または複数の空車を前記空車位置データベースから特定して、地図上に該顧客の位置と該一または複数の空車の位置とが表示される空車状況情報を作成し、該空車状況情報を該顧客端末へ送信するサーバと、を有することを特徴とする、タクシー捕捉システム。

【請求項 2】

前記基地局は、車両の属性情報を記憶する車両データベースをさらに有し、前記サーバは、前記一または複数の空車に関する属性情報を前記車両データベースから読み出して、前記空車状況情報に付加することを特徴とする、請求項 1 に記載のタクシー捕捉システム。

【請求項 3】

基地局、顧客端末、および車載装置からなるタクシー捕捉システムにおいて、前記基地局は、前記顧客端末から受信した顧客特定情報を、前記顧客端末から受信した空車選択通知に対応する空車の車両特定情報と対応させて記憶する仲介テーブルと、前記顧客端末と前記車載装置との通信を、前記仲介テーブルの対応関係にしたがって確立する仲介コネクションと、を有することを特徴とする、タクシー捕捉システム。

【請求項 4】

前記顧客端末と前記車載装置との通信は、両者間の通話による通信、または電子メールによる通信であることを特徴とする、請求項 3 に記載のタクシー捕捉システム。

【請求項 5】

前記基地局は、前記仲介テーブルにおける対応関係が継続しているあいだ一定時間ごとに、前記顧客端末から受信した顧客の位置情報を前記車載装置へ送信するとともに、前記車載装置から受信した空車の位置情報を前記顧客端末へ送信することを特徴とする、請求項 3 または 4 に記載のタクシー捕捉システム。

【請求項 6】

前記基地局は、前記顧客端末からのキャンセル通知、前記車載装置からの乗車拒否通知、または前記車載装置からの顧客下車通知のいずれかを受信すると、前記仲介テーブルに記憶した前記顧客特定情報を消去することを特徴とする、請求項 3 乃至 5 に記載のタクシー捕捉システム。

：

：

## (1) マルチ・マルチクレームの検出

### [前提]

仮想明細書の請求項 1～請求項 6 を ChatGPT4o（以下「GPT」）に与えた後、

### [プロンプト]

上記の請求項において、マルチマルチクレームが存在すれば、指摘してください。

### [回答]

上記の請求項において、「マルチマルチクレーム」が存在するかどうかを確認するためには、ある請求項が他の複数の請求項を依存先として参照しているかどうかを確認する必要があります。

「マルチマルチクレーム」とは、ある請求項が複数の独立請求項または従属請求項を基礎にして、さらにその基礎となる請求項が他の複数の請求項を基礎にしている場合を指します。

請求項を再確認すると、以下のような構造になっています：

- 請求項1は独立請求項です。
- 請求項2は請求項1に従属しています。
- 請求項3は独立請求項です。
- 請求項4は請求項3に従属しています。
- 請求項5は請求項3または4に従属しています。
- 請求項6は請求項3乃至5に従属しています。

マルチマルチクレームが存在するかを指摘するには、請求項5と請求項6が関連しています。

- **請求項5**：「請求項3または4に記載のタクシー捕捉システム」とあるため、複数の請求項（請求項3および請求項4）に依存しています。ただし、これは通常の従属形式です。
- **請求項6**：「請求項3乃至5に記載のタクシー捕捉システム」とあるため、請求項3、4、5のいずれかに依存しています。

このように、請求項5および6が、複数の請求項を依存先として指定しているため、**請求項6が「マルチマルチクレーム」に該当します。**

[結果] 正しい回答を返した。

**【書類名】 特許請求の範囲**

**【請求項 1】**

基地局、顧客端末、および車載装置からなるタクシー捕捉システムにおいて、前記基地局は、複数の空車の車載装置から受信したそれぞれの空車の位置情報を格納する空車位置データベースと、地図を記憶する地図データベースと、

前記顧客端末から受信した顧客の位置情報に基づいて、該顧客の周囲に存在する一または複数の空車を前記空車位置データベースから特定して、地図上に該顧客の位置と該一または複数の空車の位置とが表示される空車状況情報を作成し、該空車状況情報を該顧客端末へ送信するサーバと、を有することを特徴とする、タクシー捕捉システム。

**【請求項 2】**

前記基地局は、車両の属性情報を記憶する車両データベースをさらに有し、前記サーバは、前記一または複数の空車に関する属性情報を前記車両データベースから読み出して、前記空車状況情報に付加することを特徴とする、請求項 1 に記載のタクシー捕捉システム。

**【請求項 3】**

基地局、顧客端末、および車載装置からなるタクシー捕捉システムにおいて、前記基地局は、前記顧客端末から受信した顧客特定情報を、前記顧客端末から受信した空車選択通知に対応する空車の車両特定情報と対応させて記憶する仲介テーブルと、前記顧客端末と前記車載装置との通信を、前記仲介テーブルの対応関係にしたがって確立する仲介コネクションと、

を有することを特徴とする、**請求項 4 記載の**タクシー捕捉システム。



## (2) 従属関係の不備の検出

### **【前提】**

赤字・下線部のエラーを追記した特許請求の範囲を与えた後、

### **【プロンプト】**

上記の特許請求の範囲において、不適切なクレームの従属関係の記載があれば、指摘してください。

### **【回答】**

上記の特許請求の範囲において、不適切なクレームの従属関係は以下の通りです。

### **不適切な従属関係**

請求項3において、「請求項4記載のタクシー捕捉システム」と記載されています。

### **問題点**

通常、請求項はその前に記載された請求項に依存する形で記載されます。つまり、請求項3は、請求項1または請求項2に依存すべきであり、後に記載される請求項4に依存することはできません。これは、特許法の一般的な原則に反しており、不適切な記載です。

### **解決方法**

請求項3が依存すべき対象を修正し、例えば「請求項1または請求項2記載のタクシー捕捉システム」とすることで、適切な従属関係に修正することが必要です。具体的には、以下のように修正します。

### **【請求項3】**

基地局、顧客端末、および車載装置からなるタクシー捕捉システムにおいて、前記基地局は、前記顧客端末から受信した顧客特定情報を、前記顧客端末から受信した空車選択通知に対応する空車の車両特定情報と対応させて記憶する仲介テーブルと、前記顧客端末と前記車載装置との通信を、前記仲介テーブルの対応関係にしたがって確立する仲介コネクションと、

を有することを特徴とする、\*\*請求項1または2に記載の\*\*タクシー捕捉システム。

この修正により、従属関係が適切に整います。

**【結果】** 正しい回答を返した。

【書類名】 特許請求の範囲

【請求項 1】

基地局、顧客端末、および車載装置からなるタクシー捕捉システムにおいて、前記基地局は、複数の空車の車載装置から受信したそれぞれの空車の位置情報を格納する空車位置データベースと、

**前記**地図を記憶する地図データベースと、

前記顧客端末から受信した顧客の位置情報に基づいて、該顧客の周囲に存在する一または複数の空車を前記空車位置データベースから特定して、地図上に該顧客の位置と該一または複数の空車の位置とが表示される空車状況情報を作成し、該空車状況情報を該顧客端末へ送信するサーバと、

を有することを特徴とする、タクシー捕捉システム。

【請求項 2】

前記基地局は、車両の属性情報を記憶する車両データベースをさらに有し、前記サーバは、前記一または複数の空車に関する属性情報を前記車両データベースから読み出して、前記空車状況情報に付加することを特徴とする、請求項 1 に記載のタクシー捕捉システム。

【請求項 3】

基地局、顧客端末、および車載装置からなるタクシー捕捉システムにおいて、前記基地局は、前記顧客端末から受信した顧客特定情報を、前記顧客端末から受信した空車選択通知に対応する空車の車両特定情報と対応させて記憶する仲介テーブルと、

前記顧客端末と前記車載装置との通信を、前記仲介テーブルの対応関係にしたがって確立する仲介コネクションと、

を有することを特徴とする、タクシー捕捉システム。

## 2. 技術的明確性

### (1) 「前記・・・」の不備の検出

### (2) 不明確なクレームの検出

[前提] 赤字・下線部のエラーを追記した特許請求の範囲を与えた後、

[プロンプト]

上記の特許請求の範囲において、不明確である（特許法第36条第6項第2号違反）と指摘される箇所があれば、挙げてください。

[回答] 特許請求の範囲において、不明確であると指摘される可能性がある箇所について以下に挙げます。

#### 1. 地図データベースの曖昧さ

請求項1では、「地図を記憶する地図データベース」と記載されていますが、「地図」の具体的な範囲や内容が不明確です。「地図」とはどのような種類のデータを含むのか、範囲や精度などの情報が明確にされていないため、不明確であると指摘される可能性があります。

#### 2. 顧客特定情報の定義

請求項3では、「顧客特定情報」という用語が使用されていますが、この「顧客特定情報」が具体的にどのような情報を指すのか明確にされていません。例えば、顧客の名前、電話番号、位置情報など、具体的にどのデータを指すのかが不明確であるため、この点も不明確であると指摘される可能性があります。

#### 3. 空車選択通知の具体性の欠如

請求項3で用いられている「空車選択通知」という用語についても、その内容が具体的に何を指すのかが明確ではありません。空車を選択する際に送信される情報や形式が不明であり、これも不明確であると指摘される可能性があります。

#### 4. 「仲介コネクション」の不明確さ

請求項3で使用されている「仲介コネクション」という用語も、その機能や動作が具体的にどのようなものか明確に説明されていません。このコネクションがどのように確立されるのか、技術的な詳細が不明確であり、この点も不明確であると指摘される可能性があります。

まとめ

上記の不明確な用語や概念について、具体的な定義や詳細な説明を追加することで、特許法第36条第6項第2号に基づく不明確さの指摘を回避することが可能となります。

[結果]

(1) 「前記・・・」の不備の検出はできなかった。

(2) 不明確な用語の検出は、概ね出来ている。

## [プロンプト]

以下の特許請求の範囲において、「前記」または「当該」または「該」がつく要素の前に、「前記」または「当該」または「該」がない先行する要素があるかを確認してください。先行する要素がない場合、先行詞の要素無しと出力してください。

## [回答]

以下に各請求項の先行詞の有無を確認しました。

### 請求項1

- 「前記基地局」: 先行要素あり(「基地局、顧客端末、および車載装置からなるタクシー捕捉システムにおいて」)。
- 「前記空車位置データベース」: 先行要素あり(「複数の空車の車載装置から受信したそれぞれの空車の位置情報を格納する空車位置データベース」)。
- 「前記地図」: 先行要素なし(「前記地図を記憶する地図データベース」内に記載されているが、これ以前に「地図」の要素が言及されていないため)。

:

### 請求項2

- 「前記基地局」: 先行要素あり(請求項1)。
- :
- 「前記車両データベース」: 先行要素なし(「車両の属性情報を記憶する車両データベース」の前に、車両データベースの要素が言及されていないため)。
  - 「前記空車状況情報」: 先行要素あり(請求項1)。

### 請求項3

:

- 「前記仲介テーブル」: 先行要素なし(「顧客特定情報を、前記顧客端末から受信した空車選択通知に対応する空車の車両特定情報と対応させて記憶する仲介テーブル」として初めて言及)。

:

## 【前提】

仮想明細書の【発明を実施するための形態】を与えた後、

### （１）主語の無い文の検出／修正

#### 【プロンプト】

上記の【発明を実施するための形態】の中の文で、主語の無い文を検出し、主語を補った文を、以下の形式で出力して！

〔主語の無い文〕 → 〔主語を補った文〕

#### 【回答】

以下に、主語のない文を検出し、主語を補った文を示します。

1. 〔主語の無い文〕これをUターンさせて北向きへと方向を変えさせるには時間がかかると予想される。 → 〔主語を補った文〕**顧客X**は、これをUターンさせて北向きへと方向を変えさせるには時間がかかると予想される。
2. 〔主語の無い文〕**これが**仲介テーブル107へ記入される。 → 〔主語を補った文〕**空車選択通知と顧客特定情報**が仲介テーブル107へ記入される。
3. 〔主語の無い文〕これによって、顧客Xと予約車Bとの間に、対応関係が成立している。 → 〔主語を補った文〕**仲介テーブル107**によって、顧客Xと予約車Bとの間に、対応関係が成立している。
4. 〔主語の無い文〕これが仲介コネクション106を介して、顧客Xの顧客端末501へ送信され、その画面に表示される。 → 〔主語を補った文〕**予約車Bの位置情報**が仲介コネクション106を介して、顧客Xの顧客端末501へ送信され、その画面に表示される。
5. 〔主語の無い文〕これが仲介コネクション106を介して、時々刻々予約車Bの車載装置301の画面に表示される。 → 〔主語を補った文〕**顧客Xの位置情報**が仲介コネクション106を介して、時々刻々予約車Bの車載装置301の画面に表示される。

【結果】主語の補いに成功する場合もあれば、失敗する場合もある。

## (2) 複文の検出／修正

### [プロンプト]

上記の【発明を実施するための形態】の中の文で、複文を検出し、複文を2以上の単文に分割し、以下の形式で出力して！

[複文] → [分割した2以上の単文]

### [回答]

以下に複文を検出し、2つ以上の単文に分割して示します。

1. [複文] 車載装置301は、ハードウェアとしては基本的に通常のカーナビゲーションシステムであって、本システムのためのアプリケーションプログラムをインストールしたものである。 → [分割した2以上の単文] 車載装置301は、ハードウェアとしては基本的に通常のカーナビゲーションシステムである。本システムのためのアプリケーションプログラムをインストールしたものである。
2. [複文] 顧客端末501は、ハードウェアとしては一般のスマートフォンやタブレット端末であって、本システムのためのアプリケーションプログラムをインストールしたものである。 → [分割した2以上の単文] 顧客端末501は、ハードウェアとしては一般のスマートフォンやタブレット端末である。本システムのためのアプリケーションプログラムをインストールしたものである。
3. [複文] この際、顧客特定情報である顧客Xの電話番号およびeメールアドレスは、車載装置301へ通知されることはない。 → [分割した2以上の単文] この際、顧客特定情報である顧客Xの電話番号およびeメールアドレスは、車載装置301へ通知されない。
4. [複文] これによって、顧客Xと予約車Bとの間に、対応関係が成立している。 → [分割した2以上の単文] これにより、顧客Xと予約車Bとの間に対応関係が成立する。
5. [複文] 予約車Bの運転手と顧客Xとは、直接の会話によってお互いの情報を交換し合う。 → [分割した2以上の単文] 予約車Bの運転手と顧客Xは、直接の会話によって情報を交換し合う。

### [結果]

複文の検出に失敗している。また、単文に分けた場合に、主語を補うようなことはできていない。

### (3) 長文の検出／修正

#### [プロンプト]

上記の【発明を実施するための形態】の中の文で、長文を検出し、長文を2以上の文に分割する、または長文を要約し、以下の形式で出力して！

[長文] → [分割または要約した文]

#### [回答]

以下に長文を検出し、2つ以上の文に分割するか、要約して示します。

1. [長文] 基地局101の車両データベース103には、予め、本システムに登録しているタクシーの車両の属性情報が格納されている。車両の属性情報として、「ハイブリッド小型車」、「黒塗り中型車」、「7人乗りワゴン車」、「〇〇のエアロ仕様の新車」といった車両自体の特徴や、「30年間無事故無違反」、「江東区の裏道に強い」といった運転手の自己PRなどの情報が格納され、さらに、各車両の車載装置301のIPアドレスが格納されている。→ [分割または要約した文] 基地局101の車両データベース103には、登録されているタクシーの車両属性情報が格納されている。車両属性情報には、「ハイブリッド小型車」や「黒塗り中型車」などの車両特徴と、「30年間無事故無違反」などの運転手の自己PRが含まれる。また、各車両の車載装置301のIPアドレスも格納されている。
2. [長文] 顧客Xが1人で急いで大通りを北へ行きたいのであれば、画面下方の車両の属性情報に表示されているような、江東区の裏道に強い運転手（空車A）や、7人乗りワゴン車（空車C）である必要はなく、運転手に特徴が見られない小型車である空車Bで十分である。→ [分割または要約した文] 顧客Xが1人で急いで大通りを北へ行きたい場合、特定の属性を持つタクシー（例：裏道に強い運転手の空車Aや7人乗りワゴン車の空車C）を選ぶ必要はない。特徴のない小型車である空車Bで十分である。
3. [長文] 一方、もし顧客Xが同僚5人と一緒に計6人で大通りを北方向へ行きたいのであれば、7人乗りワゴン車である空車Cを呼び寄せるのがよいと判断することができる。すなわち、空車Cは、顧客Xからかなり遠くの位置にいて、かつ顧客Xから離れる方向へ走行しているので、呼び寄せるには時間がかかるが、6人が1台に乗車することができるので、運賃を節約することができるかと判断することができる。→ [分割または要約した文] 顧客Xが同僚5人と一緒に6人で移動する場合、7人乗りワゴン車の空車Cを選ぶのがよい。空車Cは遠くにあり、呼び寄せるのに時間がかかるが、6人全員が1台に乗れるため運賃を節約できる。

#### [結果]

長文を上手く簡潔化している場合と、意味が若干違ってくるような不適切な場合とがあった。

**\*補足**

**[プロンプト]**

何文字以上を長文と考えましたか？

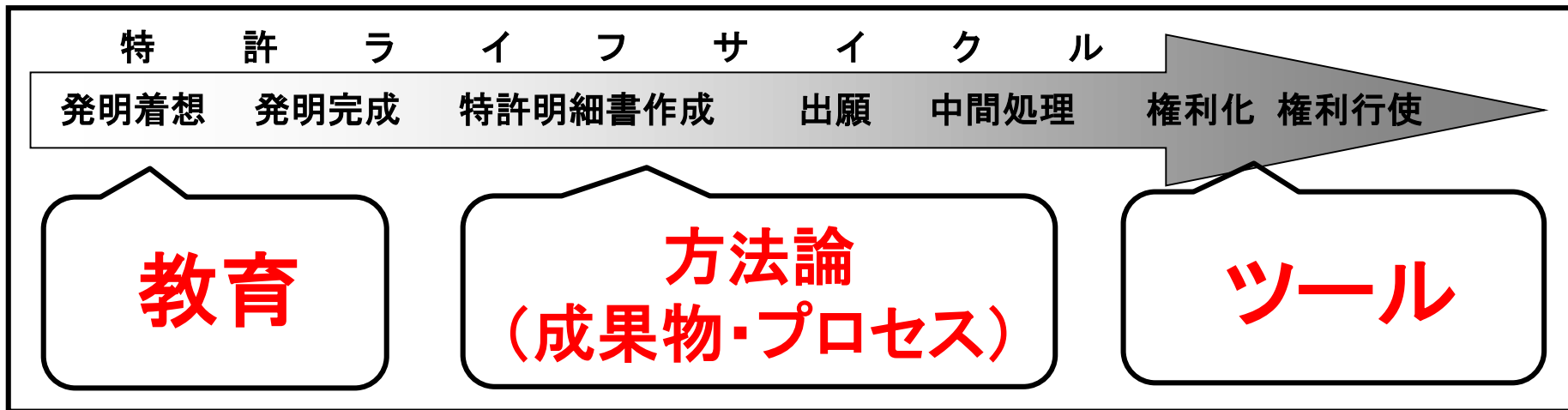
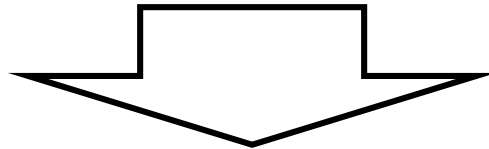
**[回答]**

今回は、約100文字以上の文章を長文として検出し、分割または要約しました。この基準は、文章の内容を分かりやすくするための目安として設定しましたが、具体的なコンテキストや文章の構造によっても異なる場合があります。



## 4. まとめ

- ・生成A I , 機械学習等のA I を用いたツールは、どんどん高性能になっていく
- ・上手くA I を使いこなす必要性が益々高まる (A I と人との協働の必要性)



弁理士業務への  
生成 A I の活用可能性と将来像  
～一人の実務者の視点から～

大澤特許事務所  
弁理士 大澤 豊

# 本日の内容

本日は、一人の実務家の視点から、以下の点についてお話しします。

- 生成 A I を、弁理士業務のどのような部分に活用し得るか。
- 生成 A I が、将来的に弁理士業務にどのような影響を与え得るか。

本日も話している内容は全て大澤個人の意見であり、所属組織とは無関係です。

# 大澤の略歴

## 大澤 豊（1974年生まれ50歳）

- 1999年 東京大学大学院農学生命科学研究科応用生命工学専攻修士課程修了
- 2000年 大澤特許事務所（東京都豊島区）入所
- 2002年 弁理士登録
- 2010年 大澤特許事務所所長

### 主な外部活動歴

- 2010年度 日本弁理士会 ソフトウェア委員会 委員長
- 2010, 2011年度 日本弁理士会 常議員
- 2011, 2012年度 実務修習講師（審査対応・演習（電気））
- 2012年度 日本弁理士会 情報企画委員会委員長
- 2012-2015, 2017年度～ 日本弁理士政治連盟 副会長
- 2015年度 日本弁理士会 執行理事
- 2016年度 日本弁理士会 副会長  
（特許委員会、バイオ・ライフサイエンス委員会等担当）
- 2017年度 日本弁理士会 経営基盤強化委員会委員長
- 2017年度～ 弁理士企業年金基金 代議員(2023年度より専務理事)
- 2020-2023年度 (一社)東京中小企業家同友会 豊島支部 政策グループ代表
- 2021年度～ 日本弁理士会 知的財産経営センター 第1事業部長（担当副センター長）

事務所での担当業務は特許（ソフト・電気・機械系）がメインです。意匠と商標は別の弁理士が担当しています。レアケースや、複数分野にまたがる業務は、複数人で共同して処理に当たります。

# 生成 A I との関わり

- 2023年4月 GPT-4の存在を知る
- 自然文による指示に応じて様々な文章を出力できる、という機能から、弁理士業務への利用可能性、さらには弁理士のコンペティターになる可能性を感じ、すぐに有料契約をして何ができるか試してみる。
- 日本弁理士会の委員会でこの試行の結果を説明する。
- 2023年夏ころから、事務所の業務への活用を開始。
- 2024年 産業構造審議会 知的財産分科会 第20回弁理士制度小委員会での議論を受けて日本弁理士会に設置された A I ツール利活用ガイドライン作成ワーキンググループの委員に選任される。現在ガイドラインを作成中。
- 随時、思いついた用途を試しつつ、生成AIの業務への利用を拡大中。

# 生成 A I の特徴

大澤の理解するところでは、（chatGPTを念頭に、一利用者として）弁理士業務への活用を考えるに当たり、考慮すべき生成 A I の特徴は以下のようなものです。

- 文章で指示を入力すると文章で回答を返してくれる。回答のフォーマットや考慮すべき論点などを指示に含めると、それに従った回答を返してくれる。
- 入力する文章はある程度長くても大丈夫（少なくとも数十ページ分くらいのデータは入力可能）。モデルによるが画像の入力も可能。
- かなりいろいろな情報を知っているが、最新の情報は知らないなので回答に反映できないことがある。（最近はウェブ検索もしてくれるので改善されていますが）
- 回答が正しくない場合はよくある。
- 回答を画像で出力できるものもあるが、フォーマットを細かく指定することはできない。
- 入力した情報が学習に使われ、他人が使う回答にその結果が反映されることがあり得る（利用規約、契約、設定等によりますが）。

# 生成 A I の特徴

また、印象レベルですが、以下のような事項は得意です。

- 指定したテーマについて概略の論点を挙げる。
- 多数のアイデアを提案する。
- いろいろな観点から検討する。
- 文章を翻訳や要約する。

逆に、以下のような事項は苦手です。

- 文献など、概念ではなく具体的な物を特定した回答をする。
- 誤りを含まない回答をする。
- 迅速に回答する。（一度の質問で欲しい答えが返ってくることは多くない）

# 生成 A I をどんなことに使えるか

ここまでに挙げた特徴を考えると

- 正しくない回答を返してくることがよくある
  - 間違った回答を自分で判別できないことには使えない
    - ≡ ある程度スキルをもっている分野なら使える
  - 生成 A I の出力をそのまま成果物としては使えない
    - ≡ 後にチェックを入れる前提の途中工程なら使える
- 一度の質問で欲しい答えが返ってくることは多くないし、迅速性も高くない
  - 自分のスキルが高いなら、自分でやった方が速いし高品質
    - ≡ 自分のスキルに自信のある分野では使う必要性が低い

私の場合、これらを踏まえて考えられる用途は、まず**翻訳**、次に**特許調査**です。



# 生成 A I をどう使うか

翻訳に使うとして、どう使うか

- 一度の質問で欲しい答えが返ってくることは多くないし、迅速性も高くない
  - 長文の翻訳には向かない。（正確性を追求するなら効率がよくない）
- 入力した情報が学習に使われ、他人が使う回答にその結果が反映されることがあり得る
  - 秘密情報の入力できない
    - ≡ 訳したい文章そのものは入力できないこともしばしばある。
- 多数のアイデアを提案する、いろいろな観点から検討するのは得意
  - 表現の選択に悩む箇所の助言を得るために使えばよいのでは？  
これなら、公知情報の入力のみでも使える。

# 翻訳への使用例（訳語選択）

**プロンプト** : ある対象を分析する際に、その対象が何であるかが「未確定」であることを英語で何といいますか？候補をいくつか挙げ、そのニュアンスの違いを説明して下さい。

**回答（ChatGPT 4oにより2024年10月に作成。一部抜粋）** :

- ▶ **1. Undetermined**
  - **意味**: 対象が「まだ決定されていない」「特定されていない」という状態を指します。
  - **ニュアンス**: まだ結論に至っておらず、現在の段階では不明であることを示します。決定するプロセスが今後進行中であることを暗示することが多いです。
  - **使用例**: "The nature of the problem remains undetermined."
  - **適用場面**: 研究や調査の初期段階で、対象に関する情報が不足している場合などに使用されます。
- ▶ **2. Uncertain**
  - **意味**: 対象に関する「確実な情報や証拠がない」「確信が持てない」という状態を指します。
  - **ニュアンス**: 決定されていないことに加え、対象に対する理解や見解そのものが不確実であることを示唆します。どちらかという、主観的なニュアンスが強いです。
  - **使用例**: "The outcome of the experiment is still uncertain."
  - **適用場面**: 結果や将来の展開が予測できない状況や、情報があっても信頼性が低い場合に使用されます。
- ▶ **3. Unidentified**
  - **意味**: 対象が「まだ識別されていない」「正体不明」であることを意味します。
  - **ニュアンス**: 対象が存在することは分かっているものの、具体的な詳細やアイデンティティが不明な状態を指します。物理的な対象や生物などに使われることが多いです。
  - **使用例**: "The object remains unidentified."
  - **適用場面**: 未確認の物体や不明な人、動物など、物理的な対象の正体が判明していない場合に使用されます。
- ▶ **4. Ambiguous**
  - **意味**: 対象が「曖昧である」「多義的である」という状態を指します。
  - **ニュアンス**: 対象が明確に定まっていないことに加えて、複数の解釈が可能である場合に使われます。解釈に幅があるという意味合いが強調されます。
  - **使用例**: "The meaning of the statement is ambiguous."
  - **適用場面**: 言葉や表現が複数の意味を持ち得る場合や、曖昧な説明によって混乱が生じている場合に使用されます。

# 翻訳への使用例（短文訳）

**プロンプト：**どの程度Xを添加すればYから受けるダメージが無視できるかに基づいて添加量を決める、を英訳してください。いくつか訳文を提案し、それぞれの書き方を選択した理由を説明してください。

**回答（ChatGPT 4oにより2024年10月に作成。一部抜粋）：**

- ▶ **1. "Determine the amount of X to be added based on how much is needed to make the damage from Y negligible."**
  - **選択理由:** この訳文は、「Yからのダメージを無視できるようにする」という具体的な目的に焦点を当てています。「negligible（無視できる）」という語を使うことで、Yのダメージを小さくすることが最終目標であることを明確に表現しています。また、「based on how much is needed」というフレーズを使うことで、必要な添加量の計算に重点が置かれています。
- ▶ **2. "Adjust the amount of X to be added to minimize the impact from Y to a negligible level."**
  - **選択理由:** この訳文は、「Xをどれだけ添加すればYの影響を最小限にできるか」に焦点を当てた表現です。「minimize（最小限にする）」という語を用いることで、Yの影響を完全に無視できるわけではなく、限りなく小さくするというニュアンスを強調しています。そのため、Yのダメージを完全にゼロにすることが難しい状況において使いやすい表現です。
- ▶ **3. "Decide the dosage of X according to the point where the damage from Y becomes insignificant."**
  - **選択理由:** 「dosage（投与量）」という語を使うことで、特に化学的な添加や薬品の使用に適した表現になります。「insignificant（重要でない）」という語を使うことで、Yのダメージが無視できるほど小さくなるという状態を、より具体的に示しています。「Decide」という動詞の選択により、最終的な判断や決定というニュアンスが強調されます。
- ▶ **4. "Set the amount of X to make the damage caused by Y negligible."**
  - **選択理由:** この表現はシンプルかつ直接的で、「Xの量を設定すること」と「Yからのダメージを無視できるレベルにすること」に焦点を当てています。簡潔でありながら目的が明確なため、特に要件や仕様を短く伝えたい場合に適しています。

# 生成 A I をどう使うか

調査に使うとして、どう使うか

- 文献など、概念ではなく具体的な物を特定した回答をするのは苦手
  - ○○について書かれた文献を挙げて、といった使い方は向かない。
  - ○○の分野の文献に付される I P C を教えて、なら多少使える。
- 多数のアイデアを提案する、いろいろな観点から検討するのは得意
  - キーワードをいろいろ提案してもらうのはあり得るのでは？

# 調査への使用例（文献抽出・失敗例）

**プロンプト:** バーコードリーダで読み取ったバーコードのデコード精度を高める技術について記載された特許公開公報をいくつか挙げてください。

**回答（ChatGPT 4oにより2024年10月に作成。一部抜粋）:**

挙げられた番号の公報の内容は、説明されているものとは全く異なります。情報の取得が難しい、と回答する場合があります。

▶ **1. 特許公開公報: JP2009145678A**

- ・ **発明の内容:** バーコードのデコード精度を高めるための新しいアルゴリズムを用いた技術。特に、バーコードが一部欠損している場合の補正技術に焦点を当てています。

▶ **2. 特許公開公報: JP2011212345A**

- ・ **発明の内容:** デコード時のノイズ除去技術を中心とした技術。外部環境からの干渉を軽減し、精度を向上させるためのフィルタリング技術を含んでいます。

▶ **3. 特許公開公報: JP2015002345A**

- ・ **発明の内容:** 高速なデコードを可能にするための新しいイメージプロセッシング技術を開発し、曖昧なバーコードでも正確に読み取ることができる技術。

▶ **4. 特許公開公報: JP2020037896A**

- ・ **発明の内容:** カメラベースのバーコードリーダ用のデコード精度を高めるためのAIを活用した技術。機械学習を用いて、読み取り精度を向上させるアプローチを取っています。

▶ **5. 特許公開公報: JP2018023456A**

- ・ **発明の内容:** 動的フォーカス調整によって、異なる距離や角度からのバーコードの正確な読み取りを実現する技術。

# 調査への使用例（IPC）

**プロンプト** : デコード処理に特徴を有するバーコードリーダに付されるIPCは何ですか？

**回答（ChatGPT 4oにより2024年10月に作成。一部抜粋）** :

デコード処理に特徴を持つバーコードリーダに関連するIPC（国際特許分類）は、通常次のような分類が用いられます：

1. **G06K 7/10** - これは、光学的な手段を用いて情報を読み取る装置に関連する分類です。特に、バーコードなどのコードを光学的に読み取る技術を含んでいます。
2. **G06K 7/12** - これは、特にコードを解釈またはデコードする技術に関する分類です。バーコードの内容を解読する際の技術や方法に関するものです。
3. **G06K 19/06** - これはバーコード自体、またはバーコードの構造や特性に関連する分類です。例えば、特定のデザインや形式のバーコードの解読に関する技術が含まれます。

これらのIPC分類は、バーコードリーダにおけるデコード処理や関連技術に対する特許を見つける際に参考となる主要な分類です。必要に応じて、より詳細な分類を探索することで、特定の技術に焦点を当てた情報が得られる可能性があります。

初めの2つは適切と考えられますが、3つ目は、記録担体に付される分類で、バーコードリーダに付されるものとしてはあまり適切でないと考えられます。また、「他にもありますか？」等として網羅性を上げようとするると、廃番になったIPCを回答してきました。

# 調査への使用例（キーワード探索）

**プロンプト：** 特許請求の範囲で「バーコードリーダー」を含む概念を指す用語として使われる語を、上位概念から下位概念までなるべく多く挙げてください。

**回答（ChatGPT 4oにより2024年10月に作成。一部抜粋）：**

## 上位概念

1. 入力装置
2. データ取得装置
3. 光学読み取り装置
4. 自動認識装置
5. 識別装置
6. スキャン装置
7. データキャプチャ装置

## 中間概念

1. コードリーダー
2. バーコード・スキャナ
3. 光学式コードリーダー
4. 二次元コードリーダー
5. マルチコードリーダー
6. ハンドヘルドリーダー
7. 固定式リーダー
8. レーザーリーダー
9. イメージングリーダー

## 下位概念

...

JplatpatのJ-GLOBALと併用することでキーワードの漏れを減らせそうです。さらに多数の候補を挙げさせることは容易ですし、その中から自分で任意に言葉を選んで検索式に入れることができます。

英語文献の検索用にはこのように英語を挙げさせることもできます。

## 上位概念

1. Data capture device
  1. データの取り込みを行う装置全般を指します。
2. Automatic identification device
  1. 自動で情報を特定する装置を広くカバーする用語です。
3. Optical scanning device
  1. 光学的なスキャン技術を用いる装置全般を指します。
4. Code reader
  4. バーコードやQRコードなどを読み取る装置を含む用語です。
5. Imaging scanner
  4. 画像取り込み機能を持つスキャナーを指します。
6. Bar code scanning system
  4. バーコードをスキャンするためのシステム全般を指します。
7. Barcode reading apparatus
  4. バーコードを読み取るための装置全般に適用される用語です。

# その他の用途

他にも、以下のような用途でいろいろ使えると思います。出力の精度は日進月歩で向上しており、現時点ではあまり役に立たないように見えても、数か月後には実用できるレベルになっている可能性もあります。

- 後で確認することを前提に情報を列挙させる
- 質を問わずに論点を多数出させる
- 正確性は別途確認する前提で長い文章を要約させる



# その他の使用例（外国O Aの要約）

本スライドの内容は投影のみとさせていただきます。

# その他の使用例（外国O Aの要約・続）

本スライドの内容は投影のみとさせていただきます。

# その他の使用例（外国O Aの要約・続）

本スライドの内容は投影のみとさせていただきます。

# その他の使用例（外国O Aの要約・続）

本スライドの内容は投影のみとさせていただきます。

# その他の使用例（商標の反論検討）

本スライドの内容は投影のみとさせていただきます。

# その他の使用例（商標の反論検討・続）

本スライドの内容は投影のみとさせていただきます。

# 生成 A I が弁理士業務に与え得る影響

- 今のところ、弁理士業務に汎用の生成 A I を活用できるシーンは限定的と考えています。
- しかし、性能の向上を考えれば、数年のうちには、活用できるシーンが相当広がるでしょう。特定の目的のためにチューニングされた生成 A I や情報の守秘性が担保される生成 A I が普及すればなおさらです。
- そうすると、社内に生成 A I を使いこなせる人材がいる企業では、現在外部の弁理士に依頼している業務の多くを社内で処理できるようになり、外部の弁理士に業務を依頼する量は大きく減るのではないのでしょうか。（生成 A I が外部弁理士の代替となるイメージ）
- 一方で、弁理士が生成 A I を活用して業務を効率化して、収益性を高められる可能性はあると考えています。
- 現在の生成 A I の立ち位置は、20年くらい前の機械翻訳と似ているのでは？

# 生成 A I が弁理士業務に与え得る影響

- 別の観点からの話として、昔は特許文献の調査は手めくりでやるしかなく、膨大な手間がかかっていました。
- 1970年代から徐々にデータベースが整備され、1999年に特許電子図書館サービスが開始されてからは、インターネット経由で誰でも無料で、手めくりより格段に少ない労力で特許文献の調査ができるようになりました。
- さて、その結果、特許情報調査の仕事は無くなったのでしょうか？
- 例えば、品質を多少犠牲にすれば生成 A I を使って今の1/3や1/10くらいの工数で明細書作成や中間処理ができるなら、それを前提とて新たな知財戦略の提案をすることも考えられるのでは。



# 生成 A I が弁理士業務に与え得る影響

「最も強い者が生き残るのではなく、  
最も賢い者が生き残るのでもなく、  
唯一、生き残る者は変化できる者である。」

～チャールズ・ダーウィン～

ご清聴  
ありがとうございました。

連絡先：〒170-0005 東京都豊島区南大塚2-33-1

ストーク南大塚 4階

大澤特許事務所

TEL: 03-5981-7885 FAX: 03-5395-3220

E-mail: [yutaka@ospat.jp](mailto:yutaka@ospat.jp)



# JaParaPat: 大規模日英特許対訳コーパス

2024/11/7(木) 第8回特許情報シンポジウム

NTT コミュニケーション科学基礎研究所 永田 昌明

## 大規模日英特許対訳コーパス JaParaPat (Japanese-English Parallel Patent Application Corpus)

- 日本と米国の特許出願データ (2000-2021) から約100万文書対の対訳文書を収集し、約3億文対の対訳コーパスを作成
- パリルートとPCTルートの両方から対訳文書を収集
- ブートストラップ: 対訳辞書を用いた文対応 → 対訳コーパス → 機械翻訳モデルを訓練 → 機械翻訳を用いた文対応 → 対訳コーパス
- 約2000万文対のWeb対訳 (JParaCrawl v3.0) に比べ、特許翻訳の精度が 約 20 BLEU ポイント向上

# 背景1: 日英の特許対訳コーパスの最新化

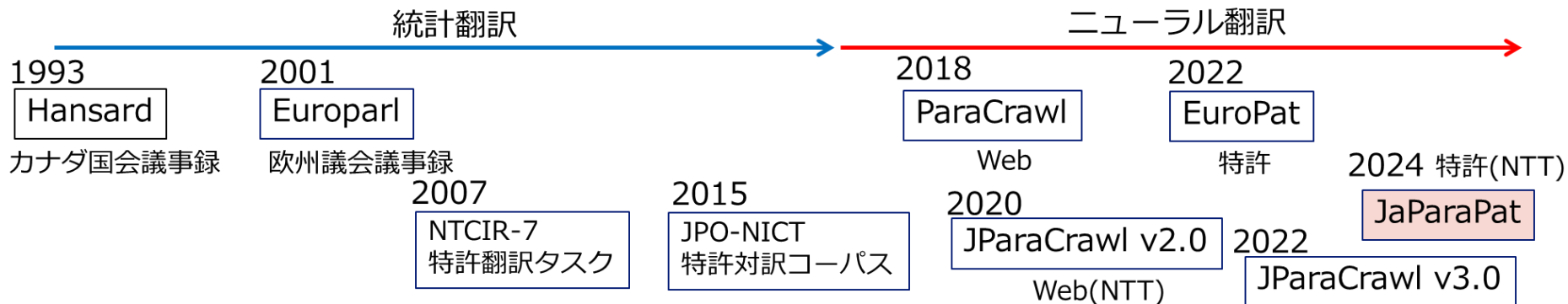


欧米では特許対訳コーパスの整備が進んでいる

- EuroPat (2022): 6つの欧州言語 (独仏西など)と英語の対訳コーパス。ParaCrawlプロジェクトの技術を応用

日本の特許対訳コーパスは内容も技術も古い

- NTCIR-7 特許翻訳タスク (2008): 約200万文対の日英対訳コーパス
- JPO・NICT英日対訳コーパス (2015): 約3.5億文対 (ALAGINから配布)



# 背景2: 特許翻訳作業の効率化

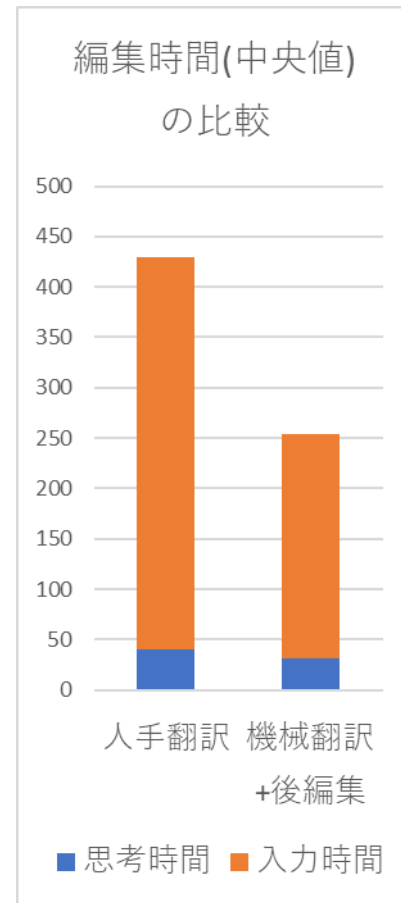
予備実験では機械翻訳の導入で作業時間を約60%に削減

- 約4000文字の原稿 4セットを、4名の翻訳者に順番を変えて提示し、以下の時間(分)を測定
  - › 編集時間: 思考時間+入力時間
  - › 思考時間: 翻訳開始から最初の編集操作まで
  - › 入力時間: 最初の編集操作から次の文の翻訳開始まで

本日の話題

特許翻訳作業の効率化のための研究開発

- 日英特許対訳コーパス [Nagata+, LREC-COLING-2024]
- 語彙制約付き機械翻訳 [Chousa and Morishita, WAT-2021]
- 編集モデルに基づく後編集支援 [Deguchi+, EAMT-2024]



# アプローチ: (J)ParaCrawl の技術の特許に適用



## JParaCrawl [Morishita+, LREC-2020/2022]

- Webから収集した日英対訳コーパス
  - › 4.8M (v1), 8.8M (v2), 21.4M (v3), 44.2M (v4)
- 基本的にはParaCrawl のコーパス構築技術を日本語化

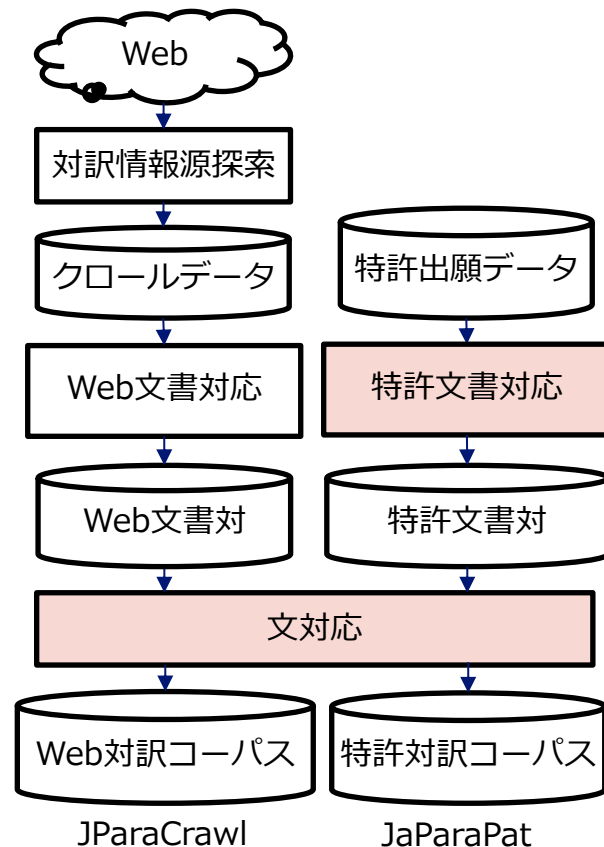
## 特許文書対応

- パリルートとPCTルートに対応した独自手法

## 文対応

- 対訳辞書に基づく手法 hunalign [Varga+, RANLP-2005] からブートストラップして、機械翻訳に基づく手法 Bleualign [Sennrich+, AMTA-2010] へ

hunalignの精度 > Bleualignの精度



## パリ ルート (パリ条約: Paris Convention)

- 最初にある国へ出願し、優先権主張して一年以内に別の国へ出願
  - › 本研究では最初に出願した国で分類: 日本 (jp-us), 米国 (us-jp), その他 (jp-x-us)

## PCT ルート (特許協力条約: Patent Cooperation Treaty)

- PCT受理官庁(receiving office)へ単一の言語と形式で出願し、30カ月以内に国内移行(national phase application)

## パテントファミリー

- 同一の発明を保護するために複数の国へ出願された特許の集合



# 一般に公開されている特許データ



## 米国特許商標庁 (United States Patent and Trademark Office, USPTO)

- 特許出願の全文データを無償で2001年3月15日から提供。対応する特許出願は一年前に日本で公開されている可能性があるため、本研究は2000年以降を対象とする

## 日本特許庁 (Japan Patent Office, JPO)

- 特許情報の一括ダウンロードサービスを無償で提供。ハードディスクを特許庁へ送る
  - › 公開特許公報: 日本語で書かれた通常の国内特許 (パリルート探索対象)
  - › 公表特許公報: 日本特許庁以外を受理官庁とするPCT国際特許を、日本へ国内移行する際に日本語へ翻訳して提出
  - › 再公開特許: 日本特許庁を受理官庁とする日本語で書かれたPCT国際特許
- 2021年12月23日に再公開特許の公開を廃止したため、本研究は2021年までを対象とする

## 欧州特許庁 (European Patent Office, EPO)

- パテントファミリーを含む全世界の特許出願の書誌データベース DOCDB を有償で提供
- 本研究では2022年4月に DOCDB を入手

# パテントファミリーに基づく 日英特許対訳文書対の抽出



網羅性の高い特許対訳文書対の探索ソフトウェアを  
独自に実装

- DOCDBのпатентファミリー情報を起点として、日本と米国で公開された特許出願の対を求める
- パリルートとPCTルートの両方を探索。日本と米国以外に最初に出願された特許も探索
- 一つの文書をタイトル、要約、本文、請求項の4つの部分に分け、各部分で文対応を計算

国際特許出願制度によって文対応の難しさが違う

- パリルート(優先権主張に基づく出願): 国によりフォーマットが異なり、必ずしも対訳ではない文書対が得られる
- PCTルート: 国際的に統一されたフォーマットが使用され、ほぼ対訳である文書対が得られる

(19) 日本国特許庁 (JP) (12) 公表特許公報 (A) (11) 特許出願公表番号  
特表2021-515539  
(P2021-515539A)

(43) 公表日 令和3年6月24日 (2021.6.24)

(51) Int. Cl.		F I	テーマコード (参考)
A 2 4 F 40/40 (2020. 01)	A 2 4 F 40/40		4 B 1 6 2
A 2 4 F 40/51 (2020. 01)	A 2 4 F 40/51		

国際出願番号 (PCTルート)

日本特許庁の特許出願公報の番号

審査請求 未請求 予備審査請求 未請求 (全 34 頁)

(21) 出願番号	特願2020-541894 (P2020-541894)	(71) 出願人	596060424
(86) (22) 出願日	平成31年3月8日 (2019. 3. 8)		フィリップ・モーリス・プロダクツ・ソシエテ・アノニム
(85) 翻訳文提出日	令和2年7月31日 (2020. 7. 31)		スイス国セアシュール2000 ヌシヤテル、ケ、ジャンルノー 3
(86) 国際出願番号	PCT/EP2019/055930	(74) 代理人	100094569
(87) 国際公開番号	W02019/170901		弁理士 田中 伸一郎
(87) 国際公開日	令和1年9月12日 (2019. 9. 12)	(74) 代理人	100103610
(31) 優先権主張番号	18161075. 9		弁理士 ▲吉▼田 和彦
(32) 優先日	平成30年3月9日 (2018. 3. 9)	(74) 代理人	100109070
(33) 優先権主張国・地域又は機関	欧州特許庁 (EP)		弁理士 須田 洋之
		(74) 代理人	100067013
			弁理士 大塚 文昭
		(74) 代理人	100086771
			弁理士 西島 幸喜

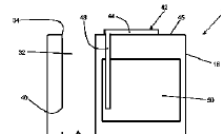
最初に欧州特許庁へ出願

最終頁に続く

(54) 【発明の名称】 カバー要素センサーを備えるエアロゾル発生装置

(57) 【要約】

ハウジング (1 2) と、エアロゾル発生物品 (8 0) を受けるためのくぼみ (3 2) と、ハウジング (1 2) によって少なくとも部分的に画定される開口部 (3 4) とを備える、エアロゾル発生装置 (1 0) が提供される。開口部 (3 4) は、開口部 (3 4) を通してくぼみ (



PCTルートの場合、同じ国際出願番号を持つ  
米国特許庁の特許出願公報を探す

# 日本特許庁 XMLファイルの例

```
<?xml version="1.0" encoding="EUC-JP"?>
<?xml-stylesheet type="text/xsl" href="../../../XSL/gat-a.xsl"?>
<!DOCTYPE jp-official-gazette PUBLIC "-//JPO//DTD PUBLISHED PATENT/UTILITY MODEL
APPLICATION 1.0//EN" "../../../DTD/gat-a.dtd">
<jp-official-gazette kind-of-jp="A" kind-of-st16="A" lang="ja" dtd-version="1.0"
country="JP" xmlns:jp="http://www.jpo.go.jp"><bibliographic-data lang="ja" coun
try="JP">
  <publication-reference>
    <document-id>
      <country>JP</country>
      <doc-number>2021093912</doc-number>
      <kind>公開特許公報(A)</kind>
      <date>20210624</date>
    </document-id>
  </publication-reference>
  <application-reference>
    <document-id>
      <doc-number>2018058673</doc-number>
      <date>20180326</date>
    </document-id>
  </application-reference>
  <invention-title>検体中に含まれる菌種を特定する方法</invention-title>
```

# 米国特許商標庁 XMLファイルの例



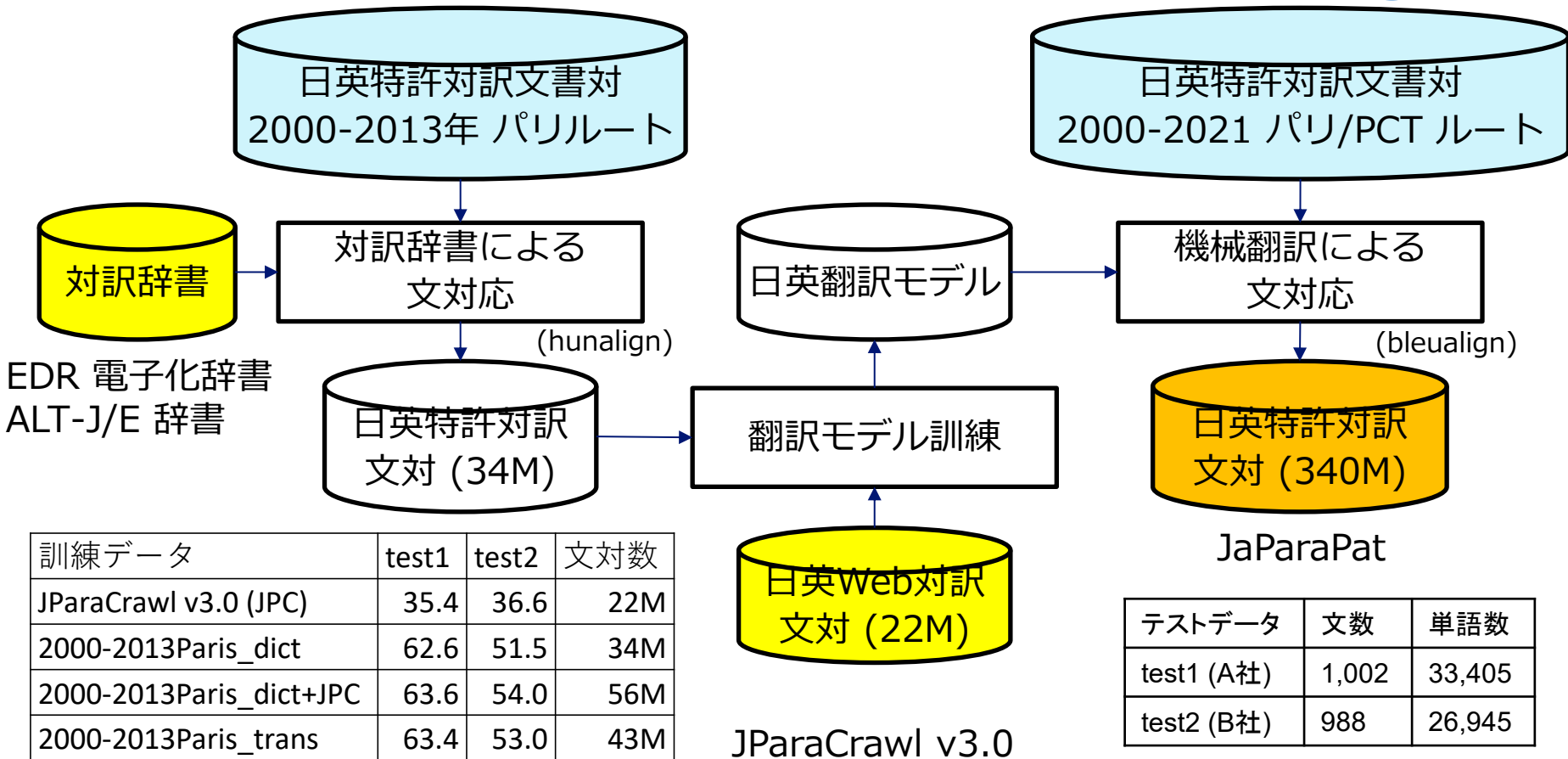
```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE us-patent-application SYSTEM "us-patent-application-v46-2022-02-17.dtd" [ ]>
<us-patent-application lang="EN" dtd-version="v4.6 2022-02-17" file="US20220295684A1-20220922.XML" status="PRODUCTION" id="us-patent-application" country="US" date-produced="20220907" date-publ="20220922">
  <us-bibliographic-data-application lang="EN" country="US">
    <publication-reference>
      <document-id>
        <country>US</country>
        <doc-number>20220295684</doc-number>
        <kind>A1</kind>
        <date>20220922</date>
      </document-id>
    </publication-reference>
    <application-reference appl-type="utility">
      <document-id>
        <country>US</country>
        <doc-number>17619810</doc-number>
        <date>20200617</date>
      </document-id>
    </application-reference>
```

## 元のXMLはもっと複雑

- Python標準ライブラリのxml.etree.ElementTreeモジュールを使って必要なデータだけを抽出

```
{
  "country": "WO",
  "kind": "A1",
  "doc-id": "550103527",
  "doc-number": "2021085030",
  "family-id": "75715054",
  "date-publ": "20210506",
  "is-representative": "YES",
  "originating-office": "EP",
  "title": "DRIVING ASSISTANCE SYSTEM",
  "publication-reference": [
    {
      "country": "WO",
      "doc_number": "2021085030"
    }
  ],
  "application-reference": [
    {
      "doc_id": "550103526",
      "country": "JP",
      "doc_number": "2020037511",
      "kind": "W"
    }
  ],
  "priority-claims": [
    {
      "doc-id": null,
      "country": "US",
      "doc_number": "201962927868",
      "kind": "P",
      "date": "20191030"
    }
  ],
  "patent-family": []
},
```

# 対訳辞書からのブートストラップ



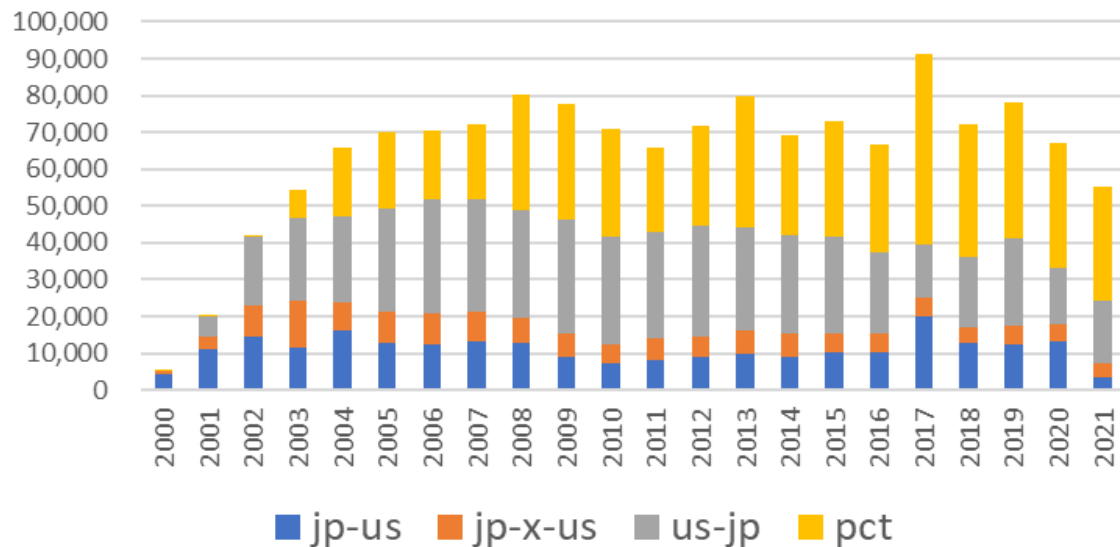
訓練データ	test1	test2	文対数
JParaCrawl v3.0 (JPC)	35.4	36.6	22M
2000-2013Paris_dict	62.6	51.5	34M
2000-2013Paris_dict+JPC	63.6	54.0	56M
2000-2013Paris_trans	63.4	53.0	43M

テストデータ	文数	単語数
test1 (A社)	1,002	33,405
test2 (B社)	988	26,945

# JaParaPat の概要



## 年別の対訳文書対数



ルート	文書数	文数	単語数
パリ	867K	182M	7.4B
PCT	527K	155M	6.2B
合計	1,394K	337M	13.6B

JP2021000998-US20210139186.ja - メモ帳

ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)

JP2021000998-US20210139186\_title\_0000\_0\_0  
JP2021000998-US20210139186\_abstract\_0000\_0\_1  
JP2021000998-US20210139186\_abstract\_0000\_1\_2  
JP2021000998-US20210139186\_abstract\_0000\_2\_3  
JP2021000998-US20210139186\_abstract\_0000\_3\_4  
JP2021000998-US20210139186\_abstract\_0000\_4\_5  
JP2021000998-US20210139186\_abstract\_0000\_5\_6  
JP2021000998-US20210139186\_abstract\_0000\_6\_7  
JP2021000998-US20210139186\_description\_0000\_0\_8  
JP2021000998-US20210139186\_description\_0001\_0\_9  
JP2021000998-US20210139186\_description\_0001\_1\_10  
JP2021000998-US20210139186\_description\_0001\_2\_11  
JP2021000998-US20210139186\_description\_0002\_0\_12  
JP2021000998-US20210139186\_description\_0003\_0\_13  
JP2021000998-US20210139186\_description\_0003\_1\_14  
JP2021000998-US20210139186\_description\_0004\_0\_15  
JP2021000998-US20210139186\_description\_0005\_0\_16  
JP2021000998-US20210139186\_description\_0006\_0\_17  
JP2021000998-US20210139186\_description\_0007\_0\_18  
JP2021000998-US20210139186\_description\_0008\_0\_19  
JP2021000998-US20210139186\_description\_0009\_0\_20  
JP2021000998-US20210139186\_description\_0010\_0\_21  
JP2021000998-US20210139186\_description\_0010\_1\_22

収納ケース

【課題】剛性を高くする。

【解決手段】収納ケース10は、前側及び上側へ開放された直方体箱状のケース。これにより、枠部材50によって、収納ケース10の前端部における剛性を高く。また、開閉パネル60が、枠部材50に回転可能に組付けられている。そして、開閉パネル60が、閉位置から前側へ回転されることで回転位置に配置。これにより、収納物を収納ケース10から出し入れするときの使用者に対する利

【選択図】図15

本発明は、収納ケースに関する。

下記特許文献1に記載の収納ケースは、上側及び前側へ開放されたボックス本体。開閉蓋部は、前面開口部を覆う閉塞位置と、天板部とボックス本体の側。そして、開閉蓋部の開放位置では、開閉蓋部が、ボックス本体の内部に。特開2016-94239号公報

しかしながら、上記収納ケースでは、天板部をボックス本体に組付けた。

このため、収納ケースの剛性を高くするという点において改善の余地が、

本発明は、上記事実を考慮して、剛性を高くすることができる収納ケー、

本発明の1又はそれ以上の実施形態は、前側及び上側へ開放された直方、

本発明の1又はそれ以上の実施形態は、前記開閉パネルの上端部には、「

本発明の1又はそれ以上の実施形態は、前記枠部材には、前記收容位置、

本発明の1又はそれ以上の実施形態は、前後方向における開閉パネルの、

本発明の1又はそれ以上の実施形態によれば、剛性を高くすることがで、

本実施の形態に係る収納ケースを示す斜視図である。

図1に示される収納ケースのスタッキング状態を示す斜視図である。



# JaParaPat: 英語文書の例



JP2021000998-US20210139186.en - メモ帳

ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)

JP2021000998-US20210139186_title_0000_0_0	STORAGE BOX
JP2021000998-US20210139186_abstract_0000_0_1	A storage box is configured to include a rectangular parallelepiped box-s
JP2021000998-US20210139186_abstract_0000_1_2	As a result, the rigidity of the front end part of the storage box can be
JP2021000998-US20210139186_abstract_0000_2_3	In addition, the opening-closing panel is rotatably assembled to the fram
JP2021000998-US20210139186_abstract_0000_3_4	The opening-closing panel is disposed at the rotation position by being r
JP2021000998-US20210139186_description_0000_0_5	The present disclosure relates to a storage box.
JP2021000998-US20210139186_description_0001_0_6	A storage box disclosed in JP-A-2016-94239 below is configured to include
JP2021000998-US20210139186_description_0001_1_7	The lid opening-closing portion is configured to be movable between a clic
JP2021000998-US20210139186_description_0001_2_8	At the open position of the lid opening-closing portion, since the lid of
JP2021000998-US20210139186_description_0002_0_9	However, in the storage box described above, when the top plate portion i
JP2021000998-US20210139186_description_0002_1_10	Therefore, there is a room for improvement in terms of increasing
JP2021000998-US20210139186_description_0003_0_11	An object of the present disclosure is to provide a storage box i
JP2021000998-US20210139186_description_0004_0_12	According to one or more embodiments of the present disclosure, a
JP2021000998-US20210139186_description_0005_0_13	In the storage box according to one or more embodiments of the pr
JP2021000998-US20210139186_description_0006_0_14	In the storage box according to one or more embodiments of the pr
JP2021000998-US20210139186_description_0007_0_15	In the storage box according to one or more embodiments of the pr
JP2021000998-US20210139186_description_0008_0_16	According to one or more embodiments of the present disclosure, t
JP2021000998-US20210139186_description_0009_0_17	FIG.1 is a perspective view illustrating a storage box according
JP2021000998-US20210139186_description_0010_0_18	FIG.2 is a perspective view illustrating a stacking state of the
JP2021000998-US20210139186_description_0011_0_19	FIG.3 is a side sectional view illustrating a state in which the
JP2021000998-US20210139186_description_0012_0_20	FIG.4 is a side sectional view (a sectional view taken along line
JP2021000998-US20210139186_description_0013_0_21	FIG.5 is a sectional view (a sectional view taken along line V-V
JP2021000998-US20210139186_description_0014_0_22	FIG.6 is a perspective view of the case main body illustrated in

日本出願番号-米国出願番号\_セクション\_段落番号\_段落内文番号\_文番号 ¥t 文

# JaParaPat: 文対応と国際特許分類の例



左側: 日本特許と米国特許の文対応

中央: 日本特許のIPC

右側: 米国特許のIPC

JP2021000998-US2...	jp-us_ipc_class_jp_2021.info - メモ帳	jp-us_ipc_class_us_2021.info - メモ帳
ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)	ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)	ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)
2 0,1	JP2021000998-US20210139186 B65D 43/20	JP2021000998-US20210139186 B65D 5/00
3 2	JP2021001262-US20210261830 C09K 3/00	JP2021001262-US20210261830 C09J 11/06
4 3	JP2021001957-US20210208529 G03G 15/20	JP2021001957-US20210208529 G03G 15/20
5 4	JP2021002617-US20210005717 H01L 21/338	JP2021002617-US20210005717 H01L 29/20
8 5	JP2021002715-US20210247578 H04B 10/80	JP2021002715-US20210247578 G02B 6/42
9 6	JP2021002959-US20210013564 B60L 58/24	JP2021002959-US20210013564 H01M 10/6563
10 7	JP2021002964-US20210009004 H02J 7/00	JP2021002964-US20210009004 B60L 58/12
11 8	JP2021002965-US20210008963 B60L 1/00	JP2021002965-US20210008963 B60H 1/00
13 9	JP2021002975-US20210006146 H02M 1/08	JP2021002975-US20210006146 H02M 1/08
14 10	JP2021003761-US20210008681 B24B 37/013	JP2021003761-US20210008681 B24B 7/24
15 11	JP2021004977-US20210364954 G03G 15/20	JP2021004977-US20210364954 G03G 15/20
16 12	JP2021005295-US20210314455 G06F 3/12	JP2021005295-US20210314455 H04N 1/00
17 13	JP2021005741-US20210274187 H04N 19/59	JP2021005741-US20210274187 H04N 19/132
18 14	JP2021005749-US20210329142 H04N 1/00	JP2021005749-US20210329142 H04N 1/00
19 15	JP2021005988-US20210296966 H02K 9/19	JP2021005988-US20210296966 H02K 9/193
20 16	JP2021006515-US20210380506 C07C 17/25	JP2021006515-US20210380506 C07C 17/087
21 17	JP2021007061-US20210357289 G11C 16/08	JP2021007061-US20210357289 G06F 11/10

# Japanese-to-English Translation Experiment

## 訓練データ

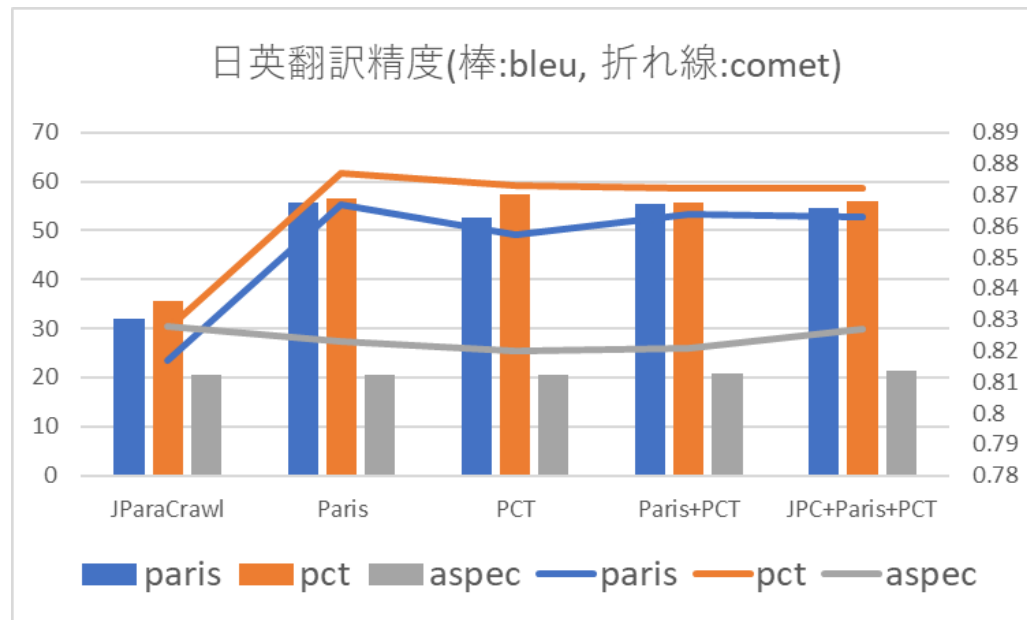
- パリ/PCTルート
  - › 2000年から2021年上半期まで
- JParaCrawl v3.0

## テストデータ

- パリ/PCTルート
  - › 2021年下半期からランダムに選択
- ASPEC (scientific paper)

## 翻訳モデル

- Transformer big



Web対訳(JParaCrawl)に比べて  
特許対訳(JaParaPat)を使うと  
翻訳精度が約20 BLEU ポイント向上

日本と米国の特許出願データから得られる日英特許対訳データの量と質を明らかにした

- 2000-2021年の出願データから約100万文書対、3億文対以上の対訳データを収集
- 55 BLEUポイント程度の翻訳精度 (Web対訳に比べ約20ポイント高い)

## 今後の課題

- 最近の特許を手で翻訳した特許翻訳のテストセットの作成
  - › 最近の出願データから収集した対訳は、ほとんどが後編集で作成されており、自動評価が混乱する
- 請求項の翻訳
  - › 請求項はとても長く、本文とは文体が異なり、特別な対処が必要
- 日中・日韓対訳データ
  - › 日本特許庁は中国や韓国の特許データも提供