Analytics Design Lab

特許文書データの分析サービス

株式会社アナリティクスデザインラボ

目次

1. 会社概要・サービス概要

企業様のデータ分析・活用を支援させていただくコンサルティング会社です

会社概要：株式会社アナリティクスデザインラボ

社名	株式会社アナリティクスデザインラボ Analytics Design Lab Inc.
所在地	〒164-0002 東京都中野区上高田1-2-51-303
設立	2017年6月1日
資本金	5,000,000円
代表取締役	野守耕爾
事業内容	・企業におけるデータ活用のコンサルティング ・新しいデータ分析技術の研究開発
取引銀行	三菱東京UFJ銀行　東中野支店
導入ツール	・Text Mining Studio（NTTデータ数理システム） ・Visual Mining Studio（NTTデータ数理システム） ・BayoLink（NTTデータ数理システム） ・SPSS Modeler（IBM） ・JMP（SAS） ・Tableau Desktop（Tableau Software）

代表略歴：野守耕爾

■ 2012年3月
早稲田大学大学院 創造理工学研究科
経営システム工学専攻　博士課程修了
博士（工学）

- 人間行動の計算モデルの開発を研究
（専門領域：人間工学）
- 2010年4月～2012年3月
独立行政法人日本学術振興会 特別研究員に採用

■ 2012年4月～（技術研修生としては2008年～）
独立行政法人産業技術総合研究所
デジタルヒューマン工学研究センター　入所

- センシング技術を応用した子どもの行動計測と人工知能技術を応用した行動の確率モデルの開発を研究

■ 2012年12月～
デロイトトーマツグループ
有限責任監査法人トーマツ
デロイトアナリティクス　入所

- データサイエンティストとしてビッグデータを活用したビジネスコンサルティング及び分析技術の研究開発に従事

■ 2017年6月～
株式会社アナリティクスデザインラボ 設立

お客様が実施されるデータ分析に対して当社が助言する「顧問アドバイザリーサービス」と、当社がデータ分析を実行しご報告する「分析プロジェクトサービス」をご用意しています

顧問アドバイザリーサービス

お客様が実施されるデータ分析・活用のお取り組みについて、専門的な観点から助言いたします

分析プロジェクトサービス

お客様の課題に応じて、データの分析と活用のプロセスを当社が設計・実行し、報告いたします

サービスの比較

お客様	分析実施者	当社
会議・メール形式での助言	提供形式	会議形式での実施報告
なし	納品成果物	報告書
3ヶ月単位	契約期間	プロジェクトの内容による

特許文書分析　相談ランキング

特許文書分析の過去のコンサルティング実績では、特許の“記述内容”に基づいて、客観的に技術を整理し、技術戦略に資する気づきを獲得したいという相談が多く寄せられます

1　客観的な技術分類

- 特許の記述内容から客観的な視点で技術を分類したい
- 自社技術と関連する技術領域の全体像を俯瞰したい

2　競合他社の動向把握

- 競合他社の特徴や棲み分け、自社との関係性を把握したい
- 他社との協業やM&Aの可能性を検討したい

3　保有技術の新規用途探索

- 自社技術を有効活用できる新しい用途を検討したい
- 自社技術を応用したイノベーションのヒントを得たい

4　事業化の技術シーズ把握

- 事業実現のための重要な技術、代替技術を把握したい
- 事業展開における競合他社の存在を把握したい

5　権利侵害リスクの把握

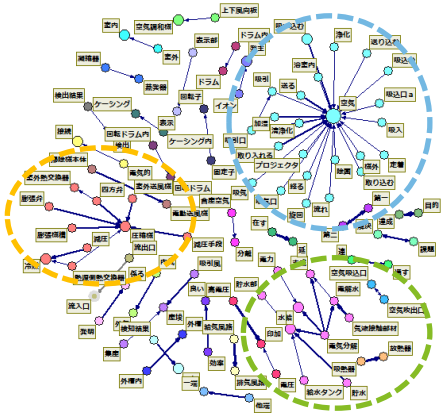
- 権利侵害になり得る類似特許を調べたい
- 従来のキーワード検索では拾えない類似特許を調べたい

2. これまでの特許文書分析の課題と解決

テキストマイニングを使ったこれまでの特許文書分析

単語をベースに、あるいは手動でグルーピングしたカテゴリをベースに、全体の出現状況、経年変化、出願人の特徴、課題と解決手段の関係などを把握する分析がよく行われます

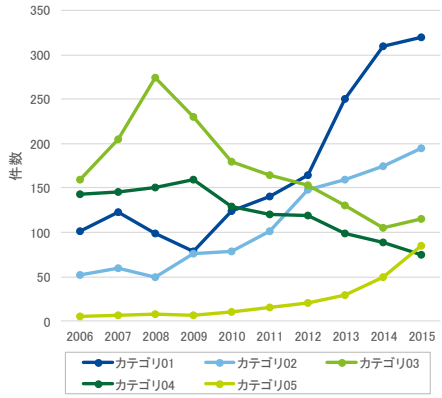
共起ネットワークによる全体像把握



- 単語の共起関係をネットワークで可視化する
- ネットワークのかたまりを見ながら、全体でどのような話題が形成されているのか考察する

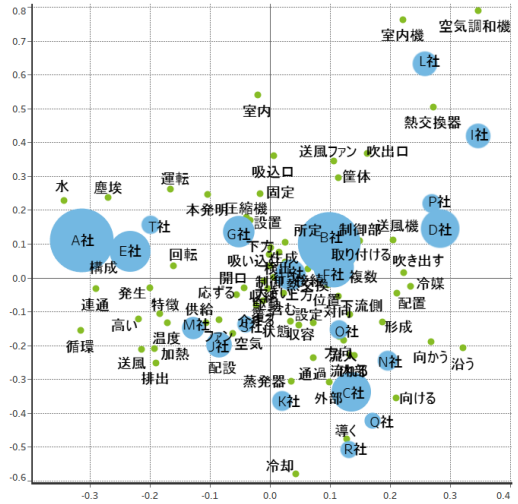
手動設定したカテゴリのトレンド把握

例）掃除機カテゴリのリスト
掃除機
集塵
集塵容器
吸引力
サイクロン
塵埃->分離
塵埃->吸い込む
塵埃->収容
塵埃->遠心分離



- 抽出した単語を手動でいくつかのカテゴリにグルーピングする
- 各カテゴリの出願年ごとの出現頻度をグラフ化し、トレンドを把握する

コレスポンデンス分析による出願人の特徴把握



- 単語の出現データから共通して現れる特徴的な軸を2つ抽出する
- その2軸による平面上に単語と出願人を同時にマッピングする
- 出願人の周辺に配置された単語群から各出願人の特徴を考察する

課題と解決手段のクロス集計による関係把握

- 「要約」の【課題】と【解決手段】それぞれに対して出現単語のカテゴリを設定する
- 課題と解決手段のカテゴリのクロス集計をして、用途と技術の関連性を考察する

これまでの特許文書分析の課題と解決アプローチ

2つのAI技術を組み合わせることで、特許文書データを単語ベースではなく、客観的に抽出されるトピックベースで解釈し、そのトピックの統計的な関連性を分析できます

課題①

単語ベースの分析では
複雑で考察しにくい

課題②

カテゴリの設定が主観的で
作業負荷も大きい

課題③

課題と解決手段の統計的な
関係を分析していない

単語を賢くクラスタリングする
AI技術

要因関係をモデリングする
AI技術

PLSA
確率的潜在意味解析

使われ方の似ている単語群を
トピックとして集約する

ベイジアンネットワーク

抽出したトピックに関わる要因
関係を統計的にモデル化する

PLSA（確率的潜在意味解析）

PLSAは、トピックモデルと呼ばれるAI技術で、テキストマイニングで抽出された大量の単語をいくつかのトピックに集約して類型化することができます

PLSAの概要

- 行列データの行の要素xと列の要素yの背後にある共通特徴となる潜在クラスzを抽出する手法である
- 元々は文書分類のための手法として開発されている(Hofman, 1999)
- 各文書の出現単語を記録した文書（行）×単語（列）という高次元（列数の多い）共起行列データに適用して複数の潜在トピックを抽出し、文書（行）×トピック（列）という低次元データに変換して文書を分類する

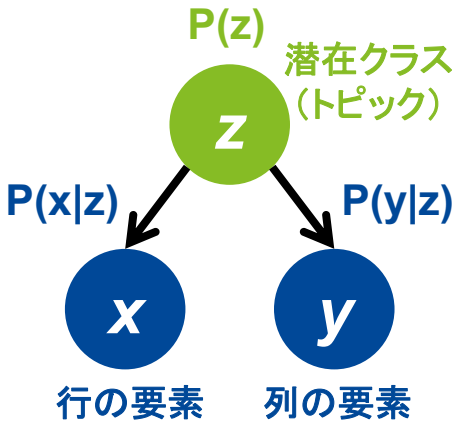
「文書×単語」行列（共起行列）

文書ID	単語1	単語2	単語3	…	単語5,014	単語5,015
1	0	0	1		1	0
2	1	0	2		0	1
…						

文書ID	トピック1	トピック2	…	トピック15
1	0.09%	0.03%		0.04%
2	0.01%	0.12%		0.06%
…				

例えば数千列ある高次元のデータでも十数個の潜在トピックで説明することができる

PLSAのグラフィカルモデル



- P(x|z), P(y|z), P(z)の3つの確率が計算される
- 潜在クラスzの数はあらかじめ設定する

※条件付確率P(A | B)
事象Bが起こる条件下で事象Aの起こる確率

xとyの共起確率を潜在クラスzを使って表現する	$P(x,y) = \sum_z P(x\|z)P(y\|z)P(z)$

PLSAのメリット

行の要素と列の要素を同時にクラスタリングできる

潜在クラスは行の要素と列の要素の2つの軸の変動量に基づいて抽出され、結果も2つの軸の情報から潜在クラスの意味を解釈することができる

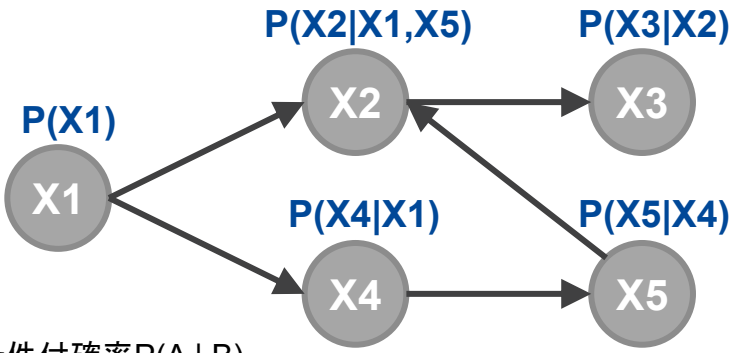
ソフトクラスタリングできる

全ての変数が全てのクラスに所属し、その各所属度合いが確率で計算されるため、複数の意味を持つ変数がある場合でも自然と表現できる

ベイジアンネットワークは、ベイズ推論に基づいたAI技術で、変数と変数の間に潜む確率的な因果関係をネットワーク構造で探索することができます

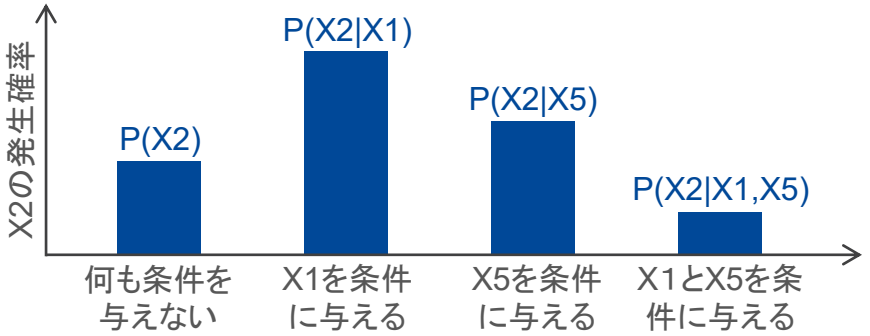
ベイジアンネットワークの概要

- 複数の変数の確率的な因果関係をネットワーク構造で表わし、ある変数の状態を条件として与えたときの他の変数の条件付確率を推論することができる
- 目的変数と説明変数の区別はなく、様々な方向から変数の確率シミュレーションができる
- 全ての変数は質的変数（カテゴリカル変数）となるため、量的変数の場合は閾値を設けてカテゴリに分割する
- 確率論の非線形処理によるモデル化のため、非線形の関係や交互作用が生じる現象でも記述できる



※条件付確率P(A | B)
事象Bが起こる条件の下で事象Aの起こる確率

確率的因果関係と交互作用



- X2の発生確率は、何も条件を与えない時（事前確率）と比べて、X1やX5を条件に与えると確率が上昇する
  ⇒X1やX5はX2の発生に関して”確率的な”因果関係がある
- しかし、X1とX5の両方を条件に与えると、元々の事前確率よりも確率が下がってしまう
  ⇒X1とX5はX2に対して交互作用がある（X1とX5は相性が悪い）

ベイジアンネットワークのメリット

現象を理解して柔軟にシミュレーションできる

目的変数、説明変数の区別なく変数の関係をモデル化するので、現象の構造を理解でき、推論変数と条件変数を自由に指定して確率推論できる

効果を発揮する有用な条件を発見できる

ある条件のときにだけ効果が現れるといった交互作用がある場合でも、確率的に意味のある関係としてモデル化することができる

AI技術の分類

AIも様々な技術があり、例えば「理解系AI」「識別系AI」「生成系AI」に分類することができますが、それぞれの技術を分析目的に応じて賢く使いこなすことが求められます

理解系AI

現状のデータに潜む特徴や要因関係を理解するAIであり、ホワイトボックスのモデルが求められる

- PLSA
- LDA
- 決定木
- ベイジアンネットワーク

「理解系AI」により、データに潜む傾向を人間が理解し、ビジネスアクションを人間が考え実行する

※私見による分類であり、一般的に定義された分類ではありません

識別系AI

画像判定や文章分類など、新規のデータを識別するAIであり、精度さえ良ければモデルはブラックボックスでもよい

- 深層学習・CNN
- RNN・LSTM・NMT
- Transformer ・T5
- BERT

生成系AI

入力した情報に対して画像や文章を生成するAIであり、精度さえ良ければモデルはブラックボックスでもよい

- GAN
- GPT
- Diffusion model

「識別系AI」や「生成系AI」は人間がデータを理解して業務に活用するのではなく、業務の自動化・省力化を目的としていることが多い

最近のAIを使った特許調査では、主に検索、要約、分類、マップ化などが行われ、その機能の本質は、AIで文書情報を複雑なベクトルに変換できたことですが、課題も見られます

AI検索

【用途】入力した文章や特許と類似度の高い特許を検索する

【仕組み】ベクトル化した特許同士の類似度を内積で計算する

【課題】ヒット理由がブラックボックスとなり妥当性を説明しにくい

AI要約

【用途】入力した特許の技術内容の要約を生成する

【仕組み】GPTに要約のプロンプトを与えて生成する

【課題】誤った要約や重要な技術的要素が省略されることがある

AIベクトル化

特許文書に対して文脈も捉えた高い表現力を持つ複雑な高次元ベクトルを学習する

AI分類

【用途】入力した特許を既定のカテゴリに自動で振り分ける

【仕組み】ベクトル化した特許と分類ラベルの間を教師あり学習する

【課題】分類数は限定的で、分類根拠はブラックボックスになる

AIマップ

【用途】特許群をその類似性に基づいて配置したマップを作成する

【仕組み】ベクトル化した特許同士の近さを計算し最適配置する

【課題】可視化結果が非常に複雑で意味解釈が難しい

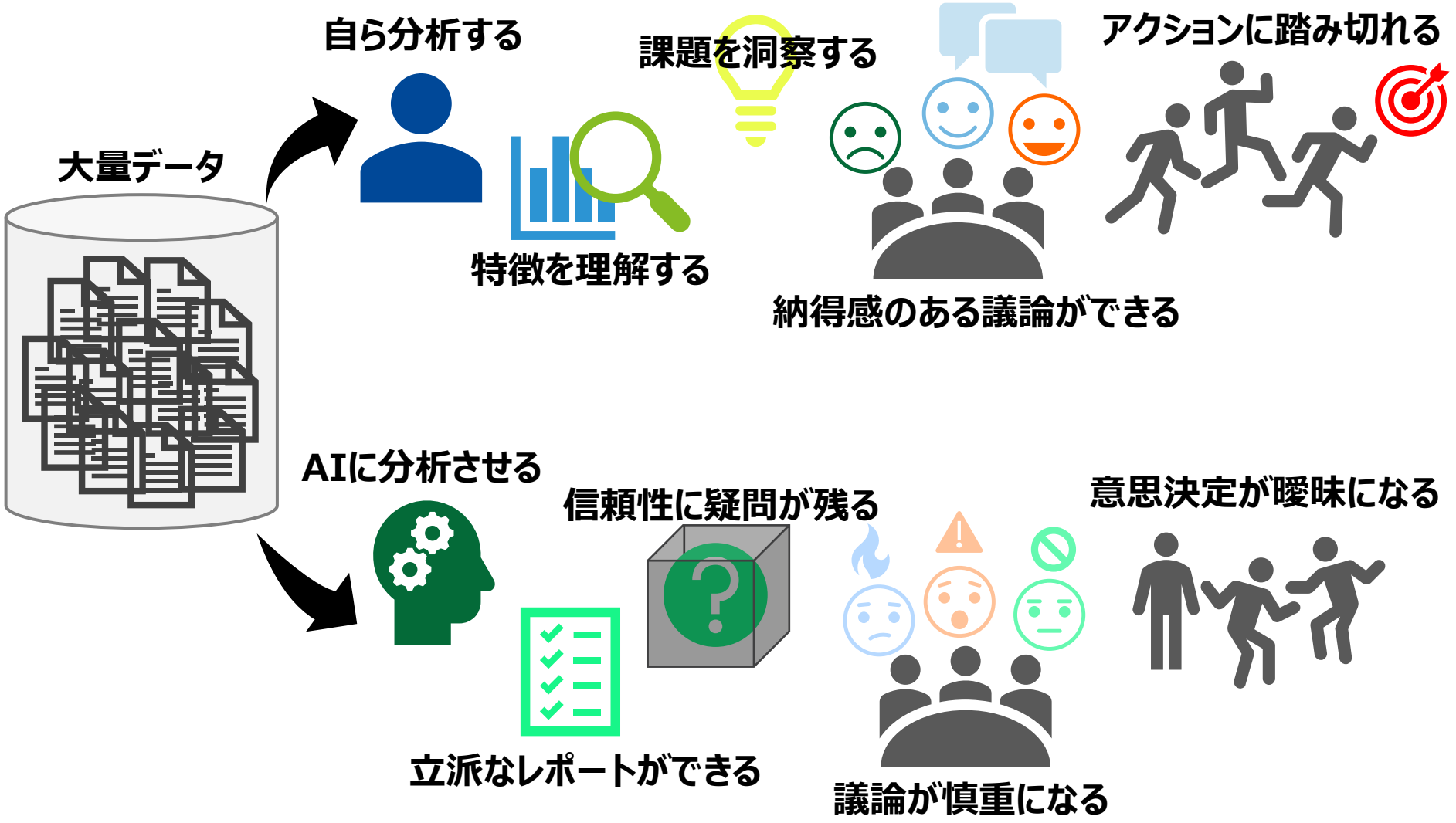
テキストデータの分析における生成系AIとテキストマイニングのメリット・デメリット

生成AIは、文脈を理解でき、汎用性が高く扱い易いですが、正確性や再現性で課題があり、テキストマイニングは、時間を要し、スキルも求められますが、正確な分析が可能です

	生成系AI	テキストマイニング
メリット 	■ 言葉による柔軟な指示を出すことができ、タスクの**汎用性**が極めて高い ■ アウトプットまでが**高速**である ■ **文脈**を捉えた分析ができる ■ 専門知識がなくても**扱いやすい**	■ 単語の出現頻度という事実ベースの分析で、**正確で客観的**な分析ができる ■ 分析は**再現可能**で、分析のプロセスもその結果も根拠が明確である ■ **定量的**な集計や統計分析に優れている ■ **大量のデータ**でも分析できる
デメリット 	■ 事前学習した汎用モデルによる推論ベースのアウトプットであり、**ハルシネーション**が起きるリスクが常にあり、特に**定量的な情報**の処理は正確でないことが多い ■ 生成プロセスが**ブラックボックス**となる ■ 同じプロンプトで回答が異なり**再現性がない** ■ **大量のデータ**を用いた分析は難しい	■ 単語の出現頻度だけに基づく単純な分析であり、**文脈の理解には限界**がある ■ 分析に多くの**時間を要し**、専門知識や経験も求められる ■ 分析の質が**分析者のスキルに依存**する

人が理解し考えることで生まれるデータ分析の価値

AIを使えばデータ分析の自動化も夢ではなさそうですが、現場で知見ある人間が自ら分析するからこそ、深い洞察が得られ、実践的で納得感のあるアクションにつながると考えます



3. 当社の独自技術のご紹介：Nomolytics

テキストマイニング × PLSA × ベイジアンネットワーク

膨大なテキストデータをトピックに変換して解釈を容易にし、テキスト情報内に潜む要因関係をモデル化して、ビジネスアクションに有用な特徴を把握可能にします

Nomolytics：Narrative Orchestration Modeling Analytics

テキストマイニング

文章に含まれる単語を抽出し、その出現頻度を集計する

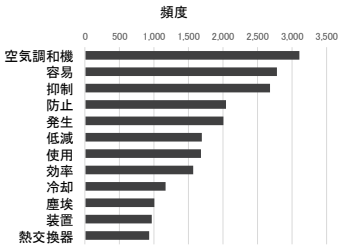
PLSA 確率的潜在意味解析

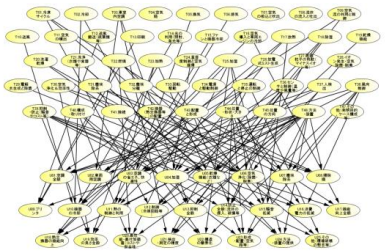
単語が出現する特徴を学習し、膨大な単語を複数のトピックにまとめる

ベイジアンネットワーク

トピックやその他属性情報など、テキスト情報内の要因関係をモデル化する

単語抽出 × **トピック化** × **因果分析**





Nomolyticsのメリット

膨大なテキストデータをいくつかのトピックという人間が理解しやすい形に整理し類型化できる

テキスト情報に潜む要因関係を構造化し、特徴を見たいターゲットのキードライバを発見できる

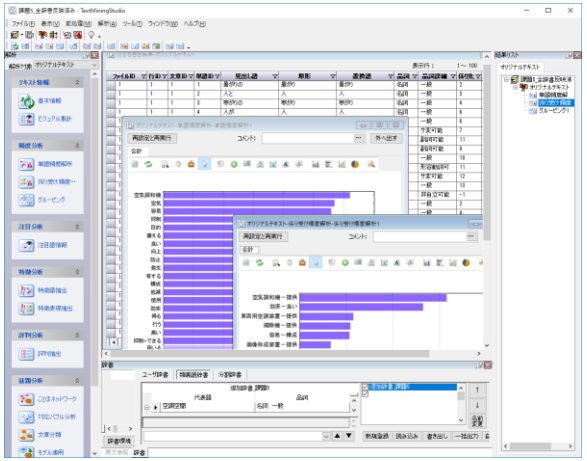
条件を変化させたときの効果を確率的にシミュレーションでき、有効なアクションを検討できる

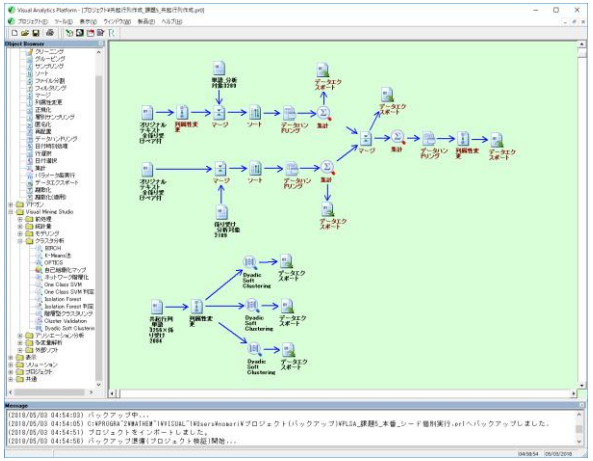
※特許登録済（特許第6085888号）

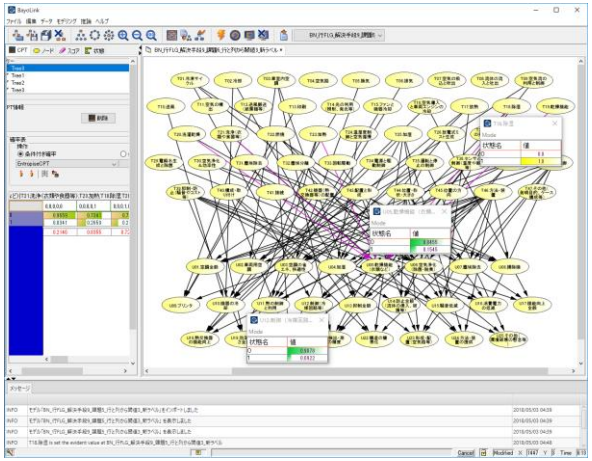
Nomolyticsを構成する各分析手法はNTTデータ数理システム社の分析ソフトウェアを使用することで実行できます

Nomolytics : Narrative Orchestration Modeling Analytics

テキストマイニング	PLSA 確率的潜在意味解析	ベイジアンネットワーク
Text Mining Studio	Visual Mining Studio	BayoLink







Nomolyticsの適用対象例

Nomolyticsは様々な業務のテキストデータに適用することができます

口コミ
- 顧客ターゲット別の関心トピックを把握
- 製品・サービス別のトピックを把握
- 口コミ得点に寄与するトピックを把握
- ニーズに応じたマーケティングを検討

アンケート
- 自由記述回答の内容をトピックで把握
- トピック化された自由記述回答と通常の定型設問回答の関係を統計分析
- 顧客満足度に寄与するトピックを把握

コールセンター履歴
- 問い合わせ内容をトピックで把握
- 製品別・顧客別のトピック傾向を把握
- 解約・退会に寄与するトピックを把握
- 満足度向上、顧客離反抑制の施策検討

特許文書
- 特許文書の内容をトピックで把握
- トレンドや競合他社の動向を把握
- 用途と技術の関係分析から用途実現の技術戦略や保有技術の新規用途を検討

営業日報
- 営業活動内容をトピックで把握
- 営業属性別のトピック傾向を把握
- 成約に寄与するトピックを把握
- 成約のための効果的な営業教育を検討

有価証券報告書
- 企業・業界の事業内容をトピックで把握
- 事業内容トピックのトレンドを把握
- 好業績に寄与する事業トピックを把握
- 定性情報から行う企業分析・業界分析

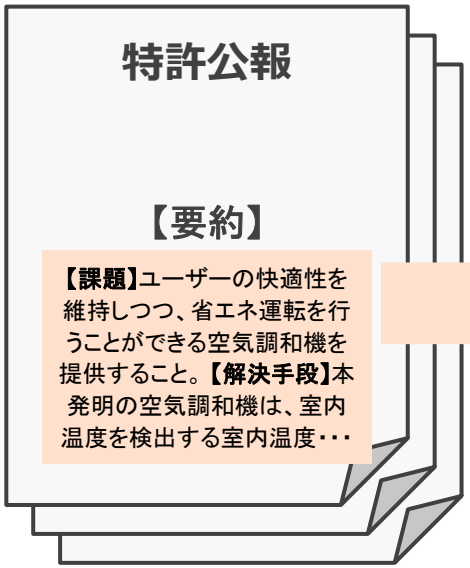
エントリーシート
- 志望動機やPR文のトピックを把握
- 記述トピックに基づいて学生を分類
- トピック傾向から面接の質問内容を検討
- 選考通過に寄与するトピックを把握

診療・看護記録
- 診療記録、看護記録をトピックで把握
- 患者の属性別のトピック傾向を把握
- 検査指標に寄与する定性情報を把握
- 定性情報も用いた診療・助言を検討

問題発生レポート
- 不具合やヒヤリハットをトピックで整理
- 作業環境別のトピック傾向を把握
- 重大問題に寄与するトピックを把握
- 問題を抑制する作業・環境改善を検討

Nomolyticsの特許文書データへの適用

特許の文書情報（要約文）にNomolyticsを適用することで、文書内容をトピックに類型化し、トピックを新たな分析の切り口として技術戦略に資する特徴を可視化できます



文書情報(要約)を分析

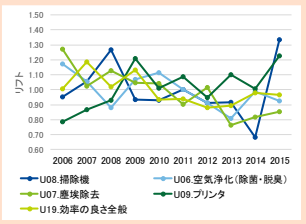
Nomolytics

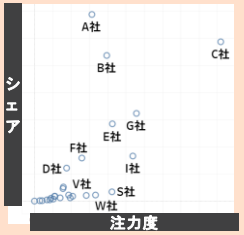
- テキストマイニング
- PLSA
- ベイジアンネットワーク

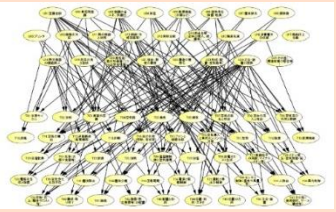
トピックに類型化して母集団の全体像を把握

トピックをベースに

トレンドを分析	競合他社の動向を分析	用途と技術の関係を分析







★★★ここがポイント★★★

従来はテキストマイニングで抽出された大量の単語をベースに分析していたが（結果が複雑だった）、それをAIで類型化されたいくつかのトピックをベースに分析することで特許文書に潜む特徴をシンプルに把握できる

4. Nomolyticsを適用した特許分析のサービス例

ある選定テーマの特許文書データに記載された内容をAI技術を応用して複数のトピックに分類・整理することで、そのテーマにおける出願特許の全体像を客観的に把握します

目的

- 選定テーマの特許の要約文に記載された内容を複数のトピックに集約する
- 要約文の中から「課題」と「解決手段」の記述を抽出し、課題からは「用途」のトピックに、解決手段からは「技術」のトピックにそれぞれ集約することも可能である

分析手順

- **(A-1)データ抽出**
  特許文書の要約文から「課題」と「解決手段」いう項目、あるいはそれに類する項目で記載されている文章を抽出する
- **(A-2)テキストマイニング**
  抽出した文章にテキストマイニングを実行し、文章に含まれる単語及びその係り受け表現(単語のペア)を抽出する
- **(A-3)辞書の作成**
  類義語辞書を作成し、辞書反映後再度テキストマイニングを実行する
- **(A-4)共起行列の作成**
  抽出結果のうち分析対象とする単語及び係り受けを選定し、「単語×単語」あるいは「単語×係り受け」の共起頻度行列を作成する
- **(A-5)PLSAの実行とトピック抽出**
  共起行列をインプットとしてPLSAを条件を調整しながら複数パターン実行し、情報量基準を目安に最適トピック解(トピック数とトピック構成語の重み)を抽出する
- **(A-6)トピック解釈**
  各トピックの構成語の重み(条件付き確率)から各トピックの意味を解釈する

期間目安　(データ件数が1～2万件程度を想定)

- 約3週間
  - 「課題」と「解決手段」からそれぞれトピックを抽出する場合は約5週間

効果

- 全ての特許文書を人が読み込まなくても全体像を把握できる
- 従来人手で行っていたカテゴリ分類を機械的に実行できるため、作業者の負荷が軽減し、客観性のある分類結果を得られる

アウトプットイメージ

B. 特許データをトピックで分類・整理する

母集団の特許データに対して各トピックのスコア(関連度)を計算し、各トピックとの関連性に応じて特許データを分類・整理し、様々な分析に活用できるデータセットを作成します

目的

- 母集団の全特許に対して抽出したトピックのスコア（関連度）を計算する

分析手順

- **(B-1)トピックのスコア計算**
  母集団の全特許データに対して、抽出したトピックのスコア（関連度）を計算する
- **(B-2)該当有無データの作成**
  計算した各トピックのスコアの閾値を設定し、トピックの該当有無のフラグ情報を付与したデータセットを作成する

期間目安　（データ件数が1～2万件程度を想定）

- 約0.5週間
  - 「課題」と「解決手段」からそれぞれ抽出したトピックについてそれぞれスコアリングする場合は約1週間

効果

- 全ての特許データを複数のトピックとそのスコアで表現できる
- 特許データと各トピックが定量的に紐づいているため、トピックに基づいて特許データを分類・整理できる
- トピックを一つの変数（属性情報）として扱えるため、他の属性情報との組み合わせで様々な分析（トレンド、企業動向、トピック間の関係性の分析など）を実行できる

アウトプットイメージ

			トピックスコア				トピックフラグ			
特許ID	出願年	出願人	トピックT01	トピックT02	トピックT03	…	トピックT01	トピックT02	トピックT03	…
1	2014	A社	2.1	1.5	5.0		0	0	1	
2	2013	B社	0.3	4.6	0.9		0	1	0	
3	2011	C社	4.8	2.7	3.1		1	0	1	
…										

C. 技術や用途のトレンドを把握する

各特許データのトピックのスコアを出願年で集計・可視化することで、技術や用途のトレンドを把握し、近年注目されるシーズやニーズを探ります

目的

■ 全特許データに対して抽出したトピックのスコア（該当度）を計算し、そのスコアを出願年で集計することで、各トピックの出願における経年変化を可視化する

分析手順

■ **（C-1）出願年とトピックの関係分析**
各出願年とトピックの関係を分析する指標値（リフト値）を計算する

■ **（C-2）トピックトレンドの可視化**
出願年を軸に各トピックの指標値（リフト値）のトレンドを可視化し、上昇傾向・下降傾向にあるトピックを抽出する

■ **（C-3）注目トレンドの特徴語の可視化**
トレンドが特徴的なトピックに該当する特許要約文にテキストマイニングを再度実行し、そのトレンドの特徴語を抽出する

期間目安　（データ件数が1～2万件程度を想定）

■ 約1週間

➢ 「課題」と「解決手段」からそれぞれ抽出したトピックについてそれぞれ分析する場合は約2週間

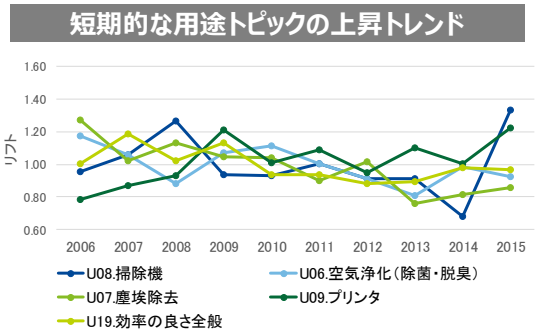
効果

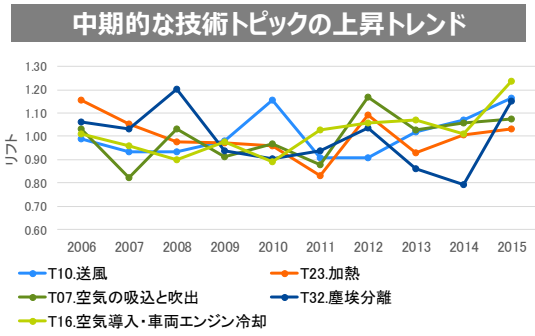
■ 上昇傾向・下降傾向にある技術や用途を把握できる

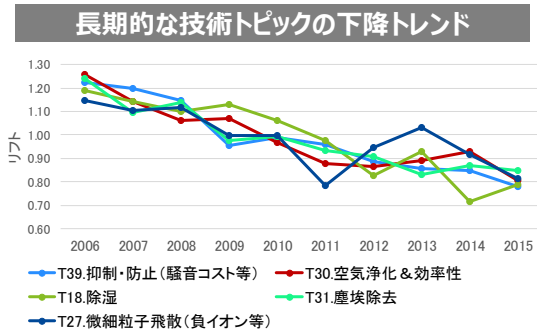
■ 用途トレンドからは、何を実現すべき技術が求められるのか察知し、その用途実現のための技術開発を検討することで、他社に先駆けて市場での優位性を得られる可能性がある

■ 技術トレンドからは、近年成長性の高そうな技術、緩やかだが長期的に堅調に推移している技術、あるいはすでに成長期が終了し飽和状態にある技術、下火になっている技術などを把握し、今後の技術の開発戦略、獲得戦略を検討できる

アウトプットイメージ







D. 競合他社の動向を把握する

各特許データのトピックのスコアを出願人で集計・可視化することで、各出願人のポジショニングや近年の出願動向を把握し、競合他社の存在や協業候補を探ります

目的

- 全特許データに対して抽出したトピックのスコア（該当度）を計算し、そのスコアを出願人で集計することで、各トピックにおける出願人の技術開発動向を可視化する

分析手順

- **（D-1）出願人情報の前処理**
  出願人情報を整理（名寄せなど）し、集計対象とする出願人を選定する
- **（D-2）出願人のシェアと注力度の計算**
  各トピックにおける出願人の動向を把握する指標として、①シェア（そのトピックの該当特許の中における各出願人の出願割合）と②注力度（各出願人の出願特許の中におけるそのトピックの該当割合）を計算する
- **（D-3）出願人の技術開発動向の可視化**
  各トピックにおいて、シェアと注力度を軸として各出願人をプロットした散布図を作成し、各社の位置づけ（静的な特徴）を把握する
- **（D-4）注目出願人の出願件数の推移の可視化**
  ポジショニングで特徴的な出願人において、そのトピックの出願件数の推移を集計し、近年での開発状況（動的な特徴）を把握する
- **（D-5）注目出願人の特徴語の可視化**
  各トピックに該当する特許要約文にテキストマイニングを再度実行し、ポジショニングで注目した出願人の出願特許の特徴語を抽出する

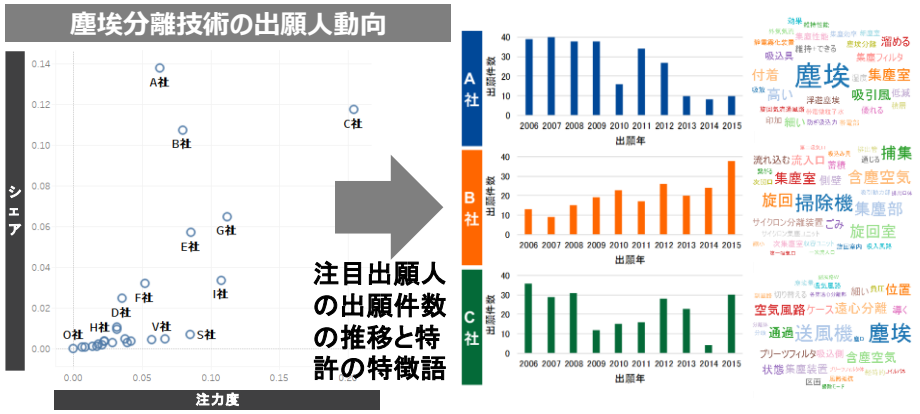
期間目安　（データ件数が1～2万件程度を想定）

- 約2週間
  - 可視化対象とするトピックは8個程度、プロットする出願人は20社程度を想定
  - 「課題」と「解決手段」からそれぞれ抽出したトピックについてそれぞれ分析する場合は約4週間

効果

- 各社のポジショニングを可視化することで、その領域における各社の位置づけを把握でき、競合となる企業や、より良いポジション獲得のために協業効果のある企業を探ることができる
- また各社の出願件数の動向や特徴語も同時に把握することで、近年の開発状況も考慮した戦略の検討ができる

アウトプットイメージ



自社と他の出願人の類似性を定量的に計算し、共同開発や技術提携、M&Aの可能性を検討します

目的

■「用途トピック」や「技術トピック」に基づいた各出願人間の類似性を定量的に計算することで、自社と類似している出願人を探索する

分析手順

■ **(E-1)出願人情報の前処理**
出願人情報を整理(名寄せなど)し、集計対象とする出願人を選定する

■ **(E-2)出願人のシェアと注力度の計算**
各トピックにおける出願人の動向を把握する指標として、①シェア(そのトピックの該当特許の中における各出願人の出願割合)と②注力度(各出願人の出願特許の中におけるそのトピックの該当割合)を計算する

■ **(E-3)出願人のクラスター分析**
各出願人の用途トピック・技術トピックの注力度(あるいはシェア)のデータにクラスター分析を実行し、それぞれの出願人間の類似性(距離)を定量的に計算する

■ **(E-4)自社と他の出願人の類似性の要因の把握**
自社と類似する他の出願人において、どのトピックが影響して類似しているのか確認する

期間目安　(データ件数が1～2万件程度を想定)

■ 約1週間
- ➢ 要約文の「課題」と「解決手段」からそれぞれ「用途トピック」と「技術トピック」を抽出していることが条件
- ➢ 類似性を確認する出願人は20社程度を想定

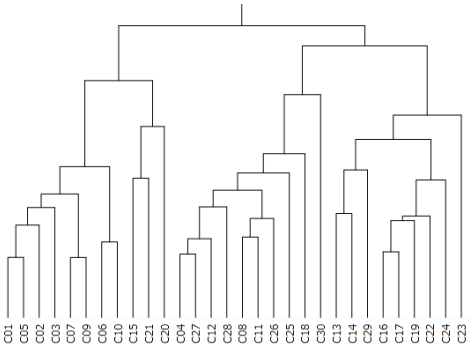
効果

■ 対象領域において、自社特許と類似している他社を効率的に探索することができる

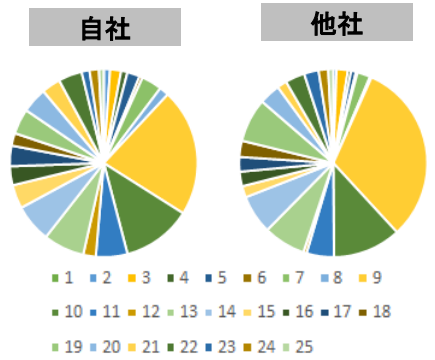
■ 対象領域の技術において、自社と相乗効果を生み出す可能性のある他社を検討し、共同開発や技術提携、あるいはM&Aといった経営戦略の意思決定を支援することができる

アウトプットイメージ

出願人に対するクラスター分析の結果



自社と類似する他の出願人のトピックの構成の比較



F. 課題を解決する技術を探索する

用途に対する技術の関係を把握することで、自社で検討中の用途実現のための重要技術と出願人動向を把握し、自社が開発・獲得すべき技術や他社との協業可能性を探ります

目的

■「用途トピック」に対する「技術トピック」の関係性を分析することで、注目用途の重要な解決技術とその出願人動向を探索する

分析手順

■ **(F-1)関係モデルの構築**
技術トピックと用途トピックのフラグデータにベイジアンネットワークを適用し、用途トピック⇒技術トピックという構造で用途と技術の確率的因果関係をモデル化する

■ **(F-2)検討用途と関係のある技術の抽出**
自社で検討中の用途トピックと統計的な関係が認められた技術トピックを抽出し、その関係度を計算する

■ **(F-3)用途×技術の出願人のシェアと注力度の計算**
検討用途と関係のある技術において、①シェア(用途×技術のトピックに該当する特許の中における各出願人の出願割合)と②注力度(各出願人の出願特許の中における用途×技術のトピックの該当割合)を計算する

■ **(F-4)用途×技術の出願人のポジショニングの可視化**
検討用途と関係のある技術において、シェアと注力度を軸として各出願人をプロットした散布図を作成し、各社の位置づけ(静的な特徴)を把握する

■ **(F-5)用途×技術での出願人の出願トレンドの可視化**
ポジショニングで特徴的な出願人において、その用途×技術のトピックの出願件数の推移を集計し、近年での開発状況(動的な特徴)を把握する

■ **(F-6)用途×技術での出願人やトレンドの特徴語の可視化**
その用途×技術のトピックに該当する特許要約文にテキストマイニングを再度実行し、注目した出願人やトレンドの特徴語やその該当原文を抽出し、その用途実現における他社の具体的な開発技術の内容を探る

期間目安　(データ件数が1～2万件程度を想定)

■ 約2週間
- ➢ 要約文の「課題」と「解決手段」からそれぞれ「用途トピック」と「技術トピック」を抽出していることが条件
- ➢ 技術を探索する用途トピックは4個程度を想定

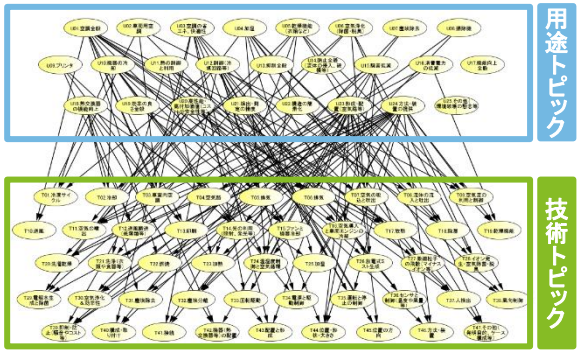
効果

■ 自社で検討中の用途に有効な技術を探索することができる

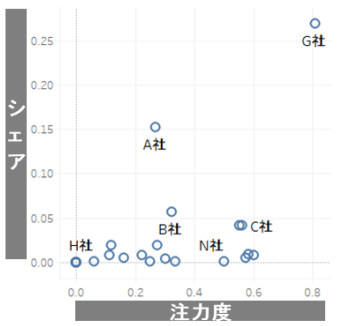
■ その有効な技術を取り巻く各社の動向を把握することで、その用途実現に向けて自社が開発あるいは獲得すべき技術、他社との直接的な競争を避け迂回手段で対抗できる代替技術、協業効果のある企業などを探索することができる

アウトプットイメージ

用途に対する技術の関係モデル



用途×技術の出願人のポジショニング



G. 技術の新しい用途を探索する

技術に対する用途の関係を把握することで、自社の技術において関係がある用途のうち、まだ想定していない用途を発見し、技術の新規用途展開のアイデアを創出します

目的

■「技術トピック」に対する「用途トピック」の関係性を分析することで、自社の保有技術の新たな用途展開を探索する

分析手順

■ **(G-1) 関係モデルの構築**
技術トピックと用途トピックのフラグデータにベイジアンネットワークを適用し、技術トピック⇒用途トピックという構造で技術と用途の確率的因果関係をモデル化する

■ **(G-2) 保有技術と関係のある用途の抽出**
自社で保有する技術トピックと統計的な関係が認められた用途トピックを抽出し、その関係度を計算する

■ **(G-3) 保有技術の新規用途候補のトピックの確認**
自社で保有する技術のトピックと統計的な関係が認められた用途トピックのうち、まだ自社で想定していない用途（該当件数が少ない用途）のトピックを保有技術の新規用途候補とする

■ **(G-4) 新規用途候補の特徴語の抽出**
自社で保有する技術トピックに該当する特許の「課題」の要約文を再度テキストマイニングし、そのなかで各新規用途候補のトピックの特徴語を抽出し、新規用途キーワードを確認する

■ **(G-5) 新規用途文の探索**
自社で保有する技術トピックに該当する特許の「解決手段」の要約文を再度テキストマイニングし、そのなかで自社の特徴的技術キーワードを抽出し、そのキーワードを「解決手段」に含む他社の特許で、新規用途候補トピックに該当する「課題」の要約文を抽出することで、自社の保有技術と関連する他社の技術の想定課題を確認し、自社の保有技術の新規用途のアイデアを具体的に検討する

期間目安　（データ件数が1～2万件程度を想定）

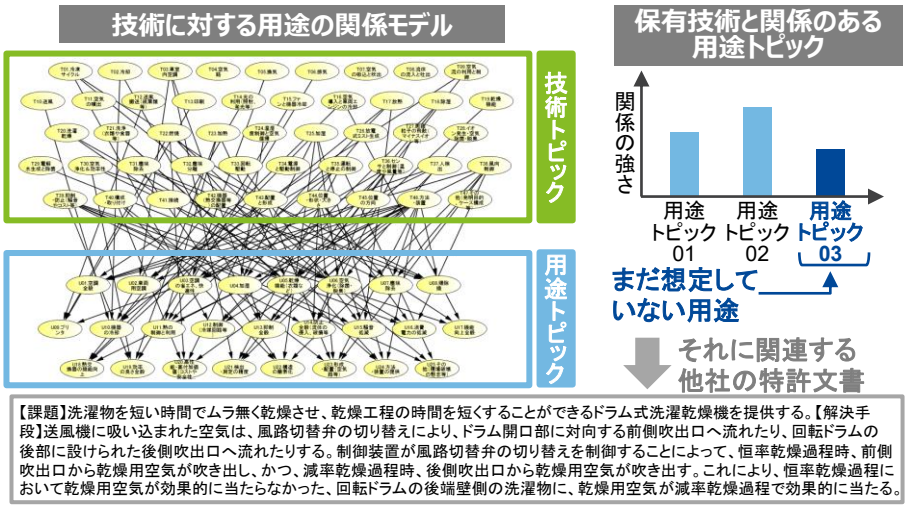
■ 約2週間
- ➢ 要約文の「課題」と「解決手段」からそれぞれ「用途トピック」と「技術トピック」を抽出していることが条件
- ➢ 用途を探索する技術トピックは4個程度を想定

効果

■ 自社の保有技術の有効な用途を探索することができる

■ まだ想定していない用途を発見し、それに関連する他の特許文書を確認することで、自社の保有技術をその用途に展開するための具体的なアイデアを創出することができる

アウトプットイメージ



【課題】洗濯物を短い時間でムラ無く乾燥させ、乾燥工程の時間を短くすることができるドラム式洗濯乾燥機を提供する。【解決手段】送風機に吸い込まれた空気は、風路切替弁の切り替えにより、ドラム開口部に対向する前側吹出口へ流れたり、回転ドラムの後部に設けられた後側吹出口へ流れたりする。制御装置が風路切替弁の切り替えを制御することによって、恒率乾燥過程時、前側吹出口から乾燥用空気が吹き出し、かつ、減率乾燥過程時、後側吹出口から乾燥用空気が吹き出す。これにより、恒率乾燥過程において乾燥用空気が効果的に当たらなかった、回転ドラムの後端壁側の洗濯物に、乾燥用空気が減率乾燥過程で効果的に当たる。

H. 課題に対する解決技術を企業別に比較する

用途に対する技術の関係を出願人別に把握し比較することで、市場での競争優位性を獲得するために効果的な提携先の検討や、強化すべき技術の検討をします

目的

■「用途トピック」に対する「技術トピック」の関係性を出願人別に分析することで、ある対象用途において各社がどのような技術を開発しているのか把握する

分析手順

■ **(H-1)出願人別モデルの構築**
技術トピックと用途トピックのフラグデータを出願人別にフィルタリングし、それぞれにベイジアンネットワークを適用し、用途トピック⇒技術トピックという構造で用途と技術の確率的因果関係を出願人別にモデル化する

■ **(H-2)用途と統計的に関係のある技術の抽出**
注目する用途トピックに対して、統計的な関係が認められた技術トピックを出願人別に抽出し、その関係度を計算する

■ **(H-3)用途と技術のクロス集計**
技術トピックと用途トピックのフラグデータをクロス集計することで、モデルでは現れなかった用途と技術の定量的関係を出願人別に計算する

■ **(H-4)用途の構成技術の出願人別比較**
注目する用途トピックに対して、各出願人がどの技術トピックの開発によってその用途を達成しようとしているのか定量的に比較する

期間目安　（データ件数が1～2万件程度を想定）

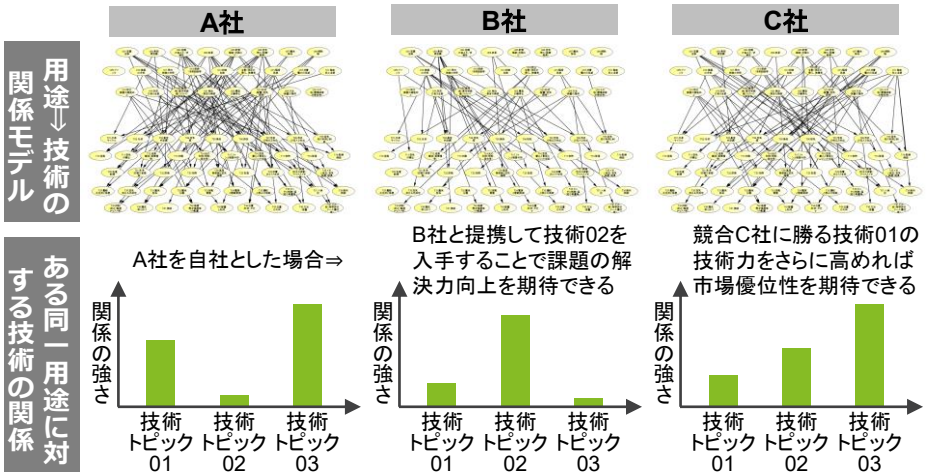
■ 約2週間
- ➢ 要約文の「課題」と「解決手段」からそれぞれ「用途トピック」と「技術トピック」を抽出していることが条件
- ➢ 対象とする出願人は自社を含めて3社程度を想定
- ➢ 分析対象とする用途トピックは3個程度を想定

効果

■ 同じ課題に対して他社の解決手段を把握できる

■ 想定課題の解決のために効果的な提携先を検討したり、競合他社に対抗するために強化すべき技術を検討できる

アウトプットイメージ



技術に対する用途の関係を出願人別に把握し比較することで、自社の技術の新たな用途の検討や、効果的な技術の売り先の検討をします

目的

■ 「技術トピック」に対する「用途トピック」の関係性を出願人別に分析することで、ある対象技術において各社がどのような用途を想定して開発しているのか把握する

分析手順

■ **(I-1)出願人別モデルの構築**
技術トピックと用途トピックのフラグデータを出願人別にフィルタリングし、それぞれにベイジアンネットワークを適用し、技術トピック⇒用途トピックという構造で技術と用途の確率的因果関係を出願人別にモデル化する

■ **(I-2)技術と統計的に関係のある用途の抽出**
注目する技術トピックに対して、統計的な関係が認められた用途トピックを出願人別に抽出し、その関係度を計算する

■ **(I-3)技術と用途のクロス集計**
技術トピックと用途トピックのフラグデータをクロス集計することで、モデルでは現れなかった技術と用途の定量的関係を出願人別に計算する

■ **(I-4)技術の用途展開の出願人別比較**
注目する技術トピックに対して、各出願人がどの用途トピックを想定して開発しているのか定量的に比較する

期間目安　（データ件数が1～2万件程度を想定）

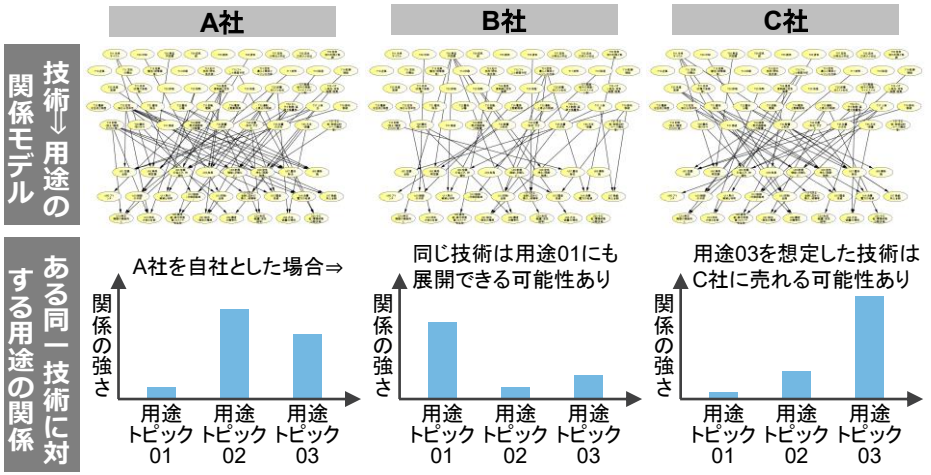
■ 約2週間
- ➢ 要約文の「課題」と「解決手段」からそれぞれ「用途トピック」と「技術トピック」を抽出していることが条件
- ➢ 対象とする出願人は自社を含めて3社程度を想定
- ➢ 分析対象とする技術トピックは3個程度を想定

効果

■ 自社と同様の技術を保有する他社の想定用途を把握できる

■ 他社にならい自社の保有技術の新たな用途を検討したり、自社の保有技術の効果的な売り先を検討できる

アウトプットイメージ



資料に関するお問い合わせやコンサルティングの
ご相談は以下までお願いします。

analytics.office@analyticsdlab.co.jp

会社ホームページもご参考にしてください。
過去の講演・論文資料や技術解説も掲載しています。

https://www.analyticsdlab.co.jp/

株式会社アナリティクスデザインラボ

Analytics Design Lab