



Analytics Design Lab

(株)技術情報協会主催セミナー  
人工知能による特許情報・技術文献調査

## AIを駆使して新たな技術戦略を導く特許文書アウェアネス

株式会社アナリティクスデザインラボ  
代表取締役 野守耕爾

2019年9月17日

## 人工知能技術を応用したデータ分析の研究開発とビジネスコンサルティングの経験を活かし、2017年6月に設立した企業様のデータ分析・活用を支援させて頂くコンサルティング会社です

### 株式会社アナリティクスデザインラボ

企業様のデータ分析・活用の支援を  
させて頂くコンサルティング会社です



データというスタートから課題の解決というゴールまでをいかにつなげばよいのか、どのようなデータ処理、分析手法、考察、アクションを検討していけばよいのか、というデータ分析を活用するプロセスを企業様の抱える課題や思惑・事情などに応じてしっかりとデザインし、それを実行することで企業様の課題解決を支援します。

設立	2017年6月1日
事業内容	● 企業におけるデータ活用のコンサルティング ● データ分析技術の研究開発
資本金	5,000,000円
所在地	東京都中野区東中野1-58-8-204
URL	<a href="http://www.analyticsdlab.co.jp/">http://www.analyticsdlab.co.jp/</a>

### 代表取締役 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

➢ 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



弊社が分析を実施しご提供する「分析受託」、お客様が実施される分析を助言する「アドバイザリー」、弊社実施の分析をお客様にトランスファーする「テラー研修」をご用意しています

### 分析受託 サービス

お客様のデータをお預かりして  
弊社がデータ分析を実施し、  
結果をご報告します

- お客様の業務課題とご提供頂くデータに応じて、弊社がデータ分析の設計を行い、実行します
- 弊社による分析の実施結果をご報告し、その報告書を成果物としてご納品します
- 分析の実施にかかる期間(作業工数)から費用をお見積りします

### アドバイザリー サービス

お客様ご自身で実施される  
データ分析・活用のご助言、  
ご指導をします

- お客様の業務課題の解決に効果的なデータ分析・活用についてご助言します
- お客様が実施される具体的なデータ分析の作業についてもご指導します
- 弊社がご納品する成果物はありません
- 1回〇時間の訪問助言を何回ご提供するのかによって費用をお見積りします

### テラー研修 サービス

弊社が実施した分析の内容を  
お客様で実施できるように、  
その手順を全てレクチャーします

- 「分析受託サービス」で弊社が実施した分析について、実施手順マニュアルや分析のプログラムファイルのご提供とともに解説し、お客様で同様の分析を実行できるように技術トランスファーします
- 「分析受託サービス」の費用に加え、マニュアルの作成や研修の実施などにかかる工数から費用をお見積りします

過去提供した特許文書分析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

## 1. 人工知能技術を応用した新たな特許文書分析

---

1-1. 従来の特許文書分析と課題

---

1-2. 人工知能技術の応用

---

1-3. 新たなテキスト分析技術: Nomolytics

---

## 2. Nomolyticsを適用した特許分析事例

---

2-1. 「風・空気」に関する特許文書データ

---

2-2. トピックの抽出

---

2-3. トピックのスコアリング

---

2-4. 出願年×トピックによるトレンド分析

---

2-5. 出願人×トピックによる競合分析

---

2-6. 用途×技術の関係分析<その1>～用途⇒技術の関係～

---

2-7. 用途×技術の関係分析<その2>～技術⇒用途の関係～

---

## 3. Nomolyticsによる特許文書分析のまとめ

---

## 4. 新たなトピックモデル技術: PCSA

---

## 5. PCSAを適用した特許分析事例

---

5-1. 「車・電気」に関する特許文書データ

---

5-2. PLSAによる全体のトピックの抽出

---

5-3. PCSAによる出願年の特徴トピックの抽出

---

5-4. PCSAによるPIスコアの特徴トピックの抽出

---

5-5. PCSAによる出願人の特徴トピックの抽出

---

## 6. まとめ

---

6-1. AIを応用した次世代テキスト分析技術  
～NomolyticsとPCSA～

---

6-2. ビッグデータからインサイト獲得のためのススメ

---

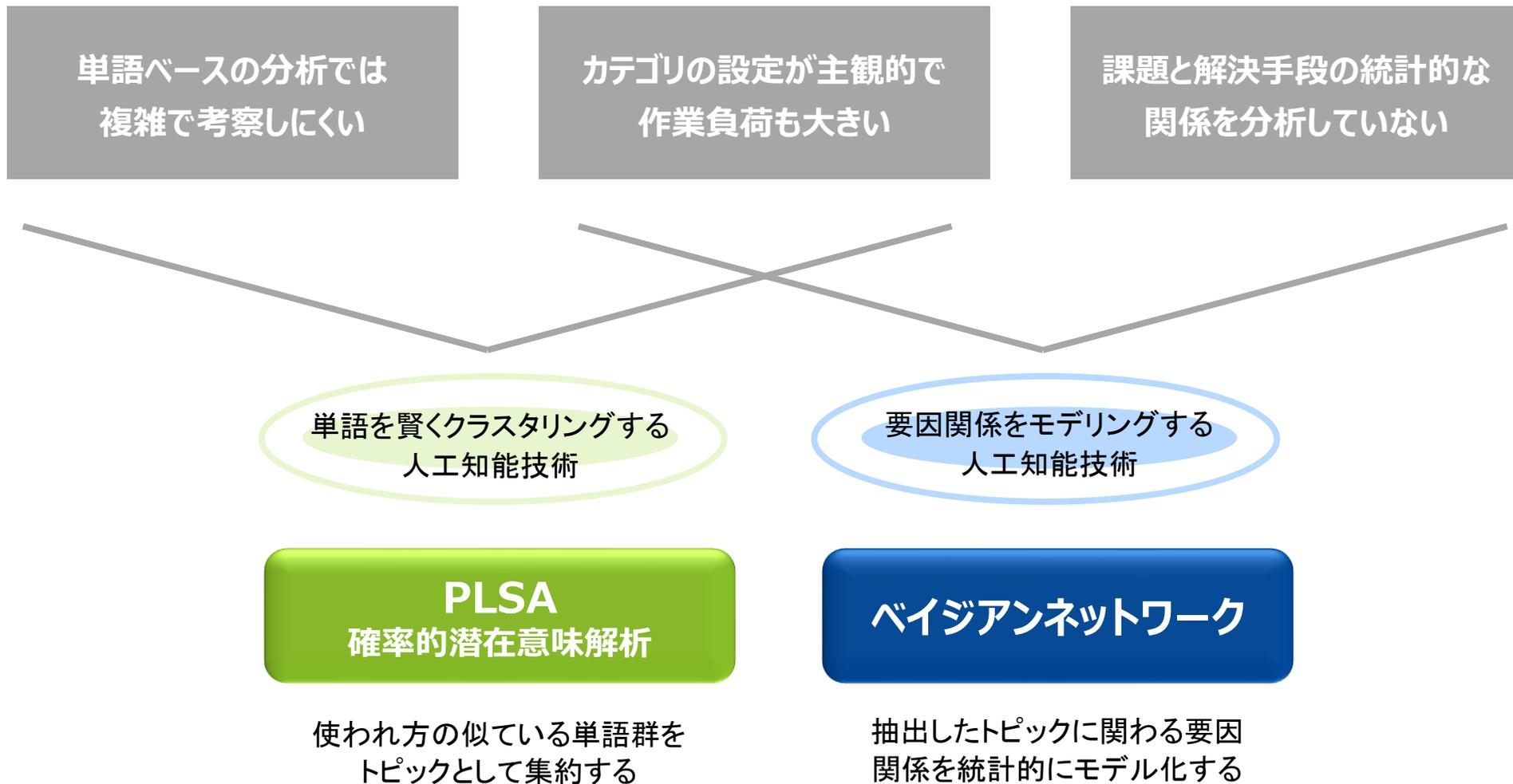
# 1. 人工知能技術を応用した新たな特許文書分析

# 1. 人工知能技術を応用した新たな特許文書分析

## 1-1. 従来の特許文書分析と課題



複数の人工知能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



# 1. 人工知能技術を応用した新たな特許文書分析

## 1-2. 人工知能技術の応用

## PLSAは、トピックモデルと呼ばれる人工知能技術で、複雑なデータをいくつかの潜在変数で説明するクラスタリング手法として用いられます

### 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

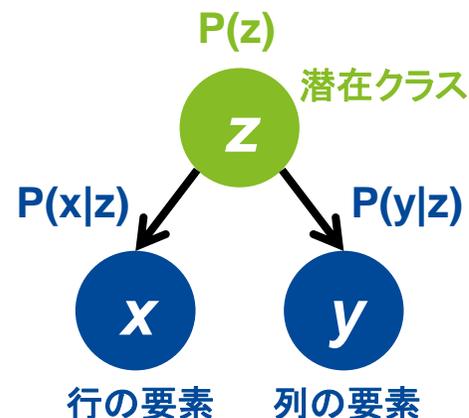
【共起行列】

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	1		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



- $P(z)$ ,  $P(x|z)$ ,  $P(y|z)$  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

$$P(x, y) = \sum_z P(z)P(x|z)P(y|z)$$

### 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

## 複雑な観測情報を分かりやすくかつ忠実に把握するため、PLSAを選択します

### 階層型 クラスタ分析

- Ward法など
- 要素間の距離を計算し、距離の近い要素同士を結合してクラスタを構成していく
- 結合の過程が樹形図で表され、結果を見てからクラスタ数を決められる(ボトムアップ的なクラスタ分析)
- データ数が多くなると計算が膨大となる

### 非階層型 クラスタ分析

- k-means法など
- あらかじめクラスタ数を決め、そのクラスタ数に全要素を一回でグルーピングする
- 各クラスタ(の重心)に対して要素の距離を計算し、距離の近い要素で集められたクラスタとなるように分類結果を調整する
- 階層型クラスタ分析よりも計算量が抑えられる

### LSA (Latent Semantic Analysis)

- 特異値分解と呼ばれる
- $(m \times n)$ の行列を、 $(m \times k), (k \times k), (k \times n)$ に分解する
- $m$ 個のデータと $n$ 個の変数を、 $k$ 個の潜在クラスで表現する(クラス数はあらかじめ設定する)
- 大きな値をとりやすいクラスが残る傾向にあるため、各要素は事前に重み付けする必要がある

### PLSA (Probabilistic Latent Semantic Analysis)

- LSAを確率的に処理
- LSAのような事前の重み付けは必要がない
- $P(x,y)$ の確率を、 $P(x|z), P(y|z), P(z)$ に分解する
- 行要素 $x$ と列要素 $y$ を、潜在クラス $z$ で表現する(クラス数はあらかじめ設定する)
- 結果は観測データのみから定義され、新規データはクラスで表現できない(過学習)

### LDA (Latent Dirichlet Allocation)

- PLSAの拡張手法
- PLSA(他左3つの手法も含め)の過学習の問題に対して、LDAではディレクレ分布を仮定し新規データのクラスを推定できる
- 新規データに対応するため、抽出されるクラスは観測データを忠実に再現するものではなく、クラスの抽象度が高い傾向がある

### 従来のクラスタ分析

- 一つの要素は必ず一つのクラスタに所属し、重複所属を許さないハードクラスタリングとなる
- 基本的に要素間の距離に基づいて分類を行う
- 要素数が多くなると要素間の距離が離れていき妥当な結果が得られにくい(次元の呪い)
- 列要素の距離に基づいて行要素を分類するか、行要素の距離に基づいて列要素を分類し、行と列どちらか一方を分類する

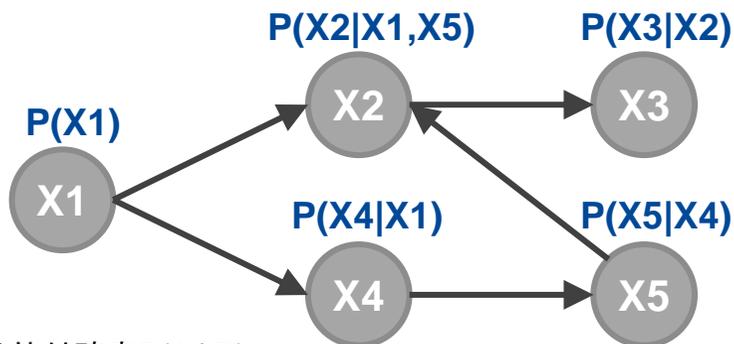
### トピックモデル

- 一つの要素は全てのクラスに所属するソフトクラスタリングで、その所属の重みを計算するため、データが複数の特徴をまたがる場合でも表現できる
- 要素間の距離の近さで分類するのではなく、高次元データの情報をできるだけ保存した形で低次元に変換する次元圧縮手法であるため、要素数が多い複雑なデータにも対応できる
- 行の要素と列の要素の背後にある共通する特徴をクラスとして抽出するため、行と列の両方をクラスタリングでき、クラスの持つ情報が多い

## ベイジアンネットワークは、ベイズ推論に基づく人工知能技術で、変数間の確率的な因果関係を探索するモデリング手法として用いられます

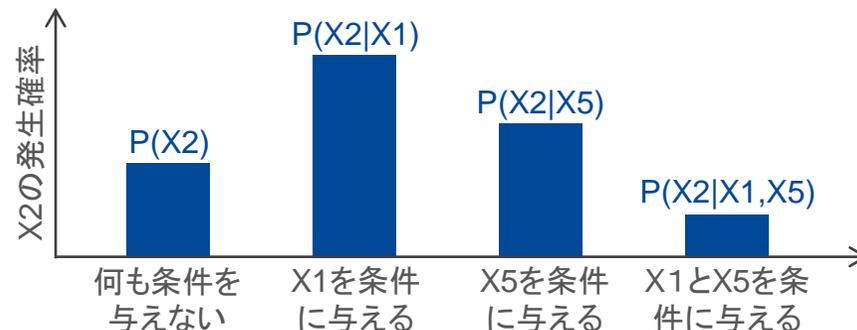
### ベイジアンネットワークの概要

- 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

## テキストで記された現象に潜む要因関係を理解するため、ベイジアンネットワークを選択します

### ニューラルネットワーク (ディープラーニング)

- 入力(説明変数)と出力(目的変数)の関係(非線形)をモデル化する
- 入力と出力の間に中間層(隠れ層)を設定し、入力情報に重みをつけて出力精度を高める処理を中間層で行う
- 柔軟性が高く複雑な関係もモデル化でき予測精度も高まるが、処理が複雑すぎてモデルの中身がブラックボックス化してしまう

### 回帰分析・判別分析 (数量化 I 類・II 類)

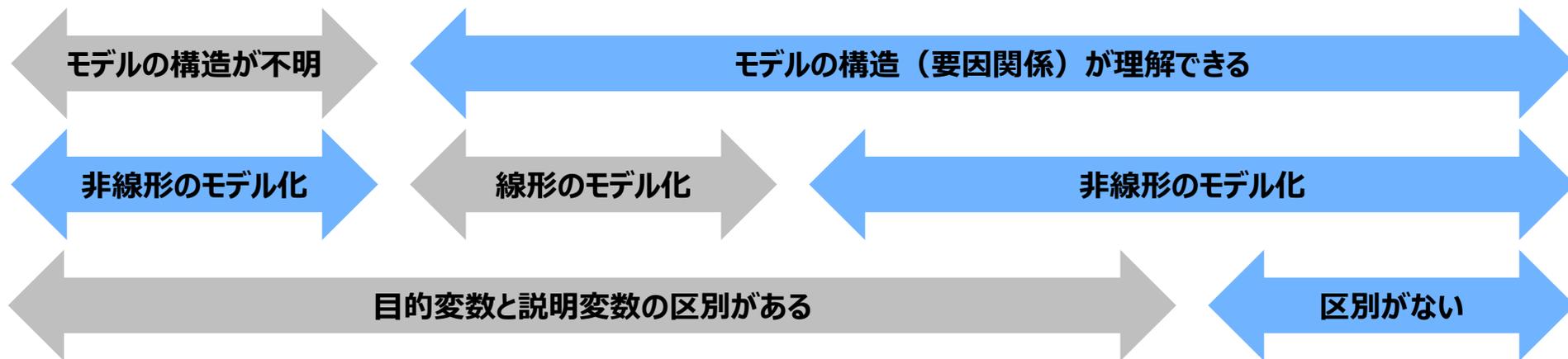
- 目的変数を説明変数の1次結合で定式化する
- 目的変数と説明変数の間に線形関係があるという仮定に基づいている
- 各説明変数の影響は独立しており、複合的な相互作用の影響は表現できない
- 説明変数間で相関が高い場合は解が不安定となり(多重共線性)、変数が多い場合この解消検討の負荷が大きい

### 決定木

- 目的変数の特徴がよく現れるルールを説明変数とその閾値による分岐で構成する
- ルールがツリー構造で可視化されるため目的変数と各説明変数の関係が分かりやすい
- 目的変数と説明変数の非線形な関係もモデル化でき、複合条件によって効果が変わる相互作用を表現しやすい

### ベイジアンネットワーク

- 複数の変数の確率的な因果関係をネットワーク構造でモデル化する
- 目的変数と説明変数の区別がないため、それぞれの変数が互いにどのような関係をもってそのデータの現象を構成しているのか理解できる
- 変数間の関係は条件付確率で計算され、複合条件によって効果が変わる相互作用も表現できる



AI技術と一言で言っても様々あり、主に新規のデータを識別予測するAIと、今あるデータが示す現象を理解するAIがあり、分析目的に応じて賢く使いこなすことが求められます

### 識別予測のAI

- ディープラーニング

- LDA

- 決定木  
(バギング、ブースティング)

新規のデータに対してそれを  
分類したり識別したりする

★ モデルの識別予測精度が命！

### 現象理解のAI

- ベイジアンネットワーク

- PLSA

- 決定木

現状のデータに潜む特徴  
や要因関係を理解する

★ モデルのホワイトボックス化が命！

# 1. 人工知能技術を応用した新たな特許文書分析

## 1-3. 新たなテキスト分析技術: Nomolytics

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

単語抽出



## 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

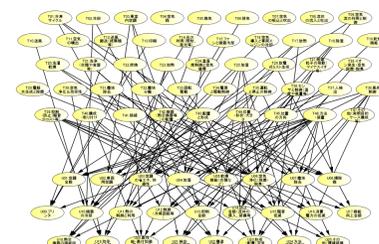
トピック抽出



## 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

モデリング



膨大なテキストデータを人間が理解しやすい形に整理できる

テキスト情報内に潜む複雑な要因関係を構造化できる

条件を変化させたときの結果の挙動をシミュレーションできる

ある事象の発生確率をコントロールする条件を発見できる



## 2. Nomolyticsを適用した特許分析事例

## 2. Nomolyticsを適用した特許分析事例

### 2-1. 「風・空気」に関する特許文書データ

## 「風」「空気」に関する10年分の特許データ30,039件の要約文を分析します

### データの抽出条件と抽出結果

- 対象
  - 公開特許公報
- キーワード
  - 要約と請求項に「風」と「空気」を含む
- 出願年
  - 2006年1月1日～2015年12月31日
- 抽出方法
  - PatentSQUAREを使用
- 抽出結果
  - 30,039件



### 分析データの加工

- 要約文の【課題】と【解決手段】に記載されている文章をそれぞれ抽出する
  - このような書式で記載されていないものは要約文をそのまま使用する
- 出願人情報は名寄せをし、グループ会社などは統一する

【要約】【課題】ユーザーの快適性を維持しつつ、省エネ運転を行うことができる空気調和機を提供すること。【解決手段】本発明の空気調和機は、室内温度を検出する室内温度検出手段と、人体の活動量を検出する人体検出手段と、基準室内設定温度を設定するリモコン装置30とを備え、室内温度が基準室内設定温度となるように空調制御を行う空気調和機であって、人体検出手段で検出する活動量が所定の活動量以内であるときは、室内温度が、基準室内設定温度を補正した補正室内設定温度となるように空調を行い、補正室内設定温度よりも低い状態を継続すると、圧縮機を停止させ、圧縮機の復帰は、基準室内設定温度に基づいて行う。

## 2. Nomolyticsを適用した特許分析事例

### 2-2. トピックの抽出

# トピック抽出のアプローチ

テキストマイニングで単語と係り受け表現を抽出し、単語×係り受けで構成される共起行列にPLSAを適用することで単語と係り受けの出現の背後にある潜在トピックを抽出します

## テキストマイニングの実行

【課題】と【解決手段】の文章に含まれる単語と係り受けを抽出する

単語	品詞	頻度
空気調和機	名詞	3,106
空気	名詞	2,846
容易	名詞	2,790
抑制	名詞	2,687
良い	形容詞	2,481
向上	名詞	2,328
防止	名詞	2,047
発生	名詞	2,005
...	...	...

係り受け表現	頻度
空気調和機⇒提供	1,575
効率⇒良い	1,325
車両用空調装置⇒提供	578
掃除機-提供	545
容易-構成	539
画像形成装置-提供	334
抑制-提供	296
向上-図る	279
...	...

## 共起行列の作成

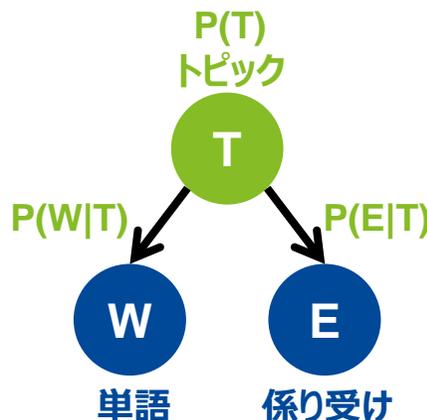
抽出した単語と係り受け表現に基づいて、「単語×係り受け」の共起行列(文章単位で同時に出現する頻度のクロス集計表)を作成する

	係り受け表現				
	空気調和機↓提供	効率↓良い	車両用空調装置↓提供	掃除機↓提供	∴
単語	空気調和機	1578	100	4	1
	空気	85	144	45	50
	容易	100	105	51	67
	抑制	142	95	64	63
	...				

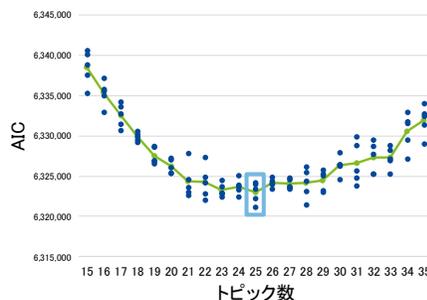
共起行列の構成(それぞれ頻度10件以上を対象)  
 課題: 単語(3,256語) × 係り受け(2,084表現)  
 解決手段: 単語(5,187語) × 係り受け(7,174表現)

## PLSAの実行

共起行列にPLSAを適用する



トピック数を幅を持たせて設定し、各トピック数に対してPLSAを初期値を変えて5回ずつ実行して情報量基準AICを計算し、AIC最小の解を採用する



## トピックの抽出

各トピックについて以下の3つの確率が計算される

- ①  $P(T)$   
トピックの存在確率
- ②  $P(W|T)$   
トピックにおける単語の所属確率
- ③  $P(E|T)$   
トピックにおける係り受けの所属確率

トピックにおける $P(W|T)$ と $P(E|T)$ からトピックの意味を解釈する

U04		
P(U)=2.6%		
P(W U)	単語	P(E U)
5.5%	加湿装置	6.8%
3.7%	水	3.1%
3.3%	供給	2.9%
2.4%	加湿	1.9%
2.3%	カビ	1.7%
2.1%	加湿器	1.5%
2.1%	発生	1.4%
2.0%	繁殖	1.3%
1.9%	ミスト	1.3%
1.7%	加湿性能	1.2%
1.5%	ミスト発生装置	1.2%
1.4%	細菌	1.2%
1.3%	室内	1.2%
1.3%	抑制+できる	1.1%
1.1%	浴室	1.1%
...	...	...

# 用途ピック25個の一覧

【課題】の文章からは、空調や加湿、空気清浄、掃除機、プリンタ、機器冷却、騒音や消費電力の低減、構造の簡素化などの用途が25個抽出されました

<h3>U01.空調全般</h3> <p>空気調和機</p>	<h3>U02.車両用空調</h3> <p>車両用空調装置</p>	<h3>U03.空調の省エネ、快適性</h3> <p>空調システム</p>	<h3>U04.加湿</h3> <p>加湿装置</p>	<h3>U05.乾燥機能 (衣類など)</h3> <p>乾燥機</p>
<h3>U06.空気浄化 (除菌・消臭)</h3> <p>空気清浄機</p>	<h3>U07.塵埃除去</h3> <p>塵埃除去</p>	<h3>U08.掃除機</h3> <p>掃除機</p>	<h3>U09.プリンタ</h3> <p>画像形成装置</p>	<h3>U10.機器の冷却</h3> <p>冷却装置</p>
<h3>U11.熱の制御と利用</h3> <p>熱制御</p>	<h3>U12.制御 (冷媒回路等)</h3> <p>冷媒制御</p>	<h3>U13.抑制全般</h3> <p>抑制</p>	<h3>U14.防止全般 (流体の侵入、破損等)</h3> <p>防止</p>	<h3>U15.騒音低減</h3> <p>騒音低減</p>
<h3>U16.消費電力の低減</h3> <p>消費電力低減</p>	<h3>U17.機能向上全般</h3> <p>機能向上</p>	<h3>U18.熱交換器の機能向上</h3> <p>熱交換器</p>	<h3>U19.効率の良さ全般</h3> <p>効率</p>	<h3>U20.価値 (コストや安全性など)</h3> <p>価値</p>
<h3>U21.検出・測定 の精度</h3> <p>検出精度</p>	<h3>U22.構造の簡素化</h3> <p>構造簡素化</p>	<h3>U23.形成・配置 (空気路等)</h3> <p>形成配置</p>	<h3>U24.方法・装置の提供</h3> <p>方法装置</p>	<h3>U25.その他 (環境破壊の懸念等)</h3> <p>その他</p>





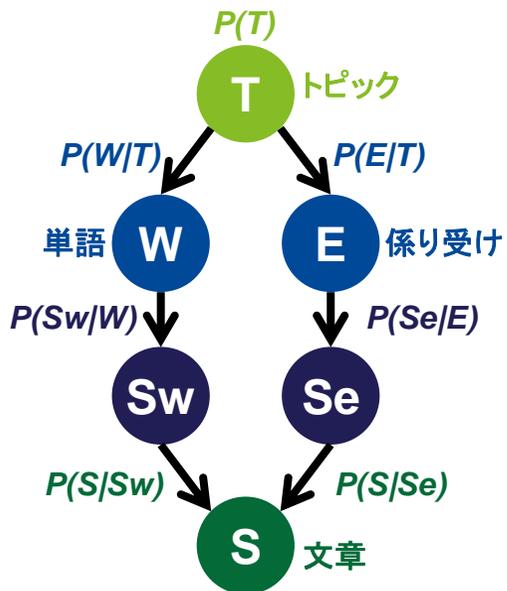
## 2. Nomolyticsを適用した特許分析事例

### 2-3. トピックのスコアリング

文章単位に各トピックのスコア(該当度)を計算し、それを特許ID単位に集約し、最終的には閾値を設定して{0:該当無,1:該当有}のデータに変換します

文章単位 のスコア	$\frac{P(S T)}{P(S)}$
--------------	-----------------------

- リフト値(事後確率÷事前確率)
- トピックを条件とすることで文章の発生確率が何倍になるのかを示す



文章を単語で定義される文章Swと係り受けで定義される文章Seを設定し、それぞれトピックとの関係を計算し、最終的にそれらを一つに統合する

単語 $W_i$ で定義される文章 $Sw_h$
$Sw_h = \{W_1, W_2, \dots, W_i\}$
トピック $T_k$ を条件とした文章 $Sw_h$ の出現確率
$P(Sw_h T_k) = \sum_i P(Sw_h W_i)P(W_i T_k)$
単語 $W_i$ が出現する中で文章 $Sw_h$ が出現する確率( $W_i$ の出現文章数の逆数)
$P(Sw_h W_i) = 1/n(W_i)$
係り受け $E_j$ で定義される文章 $Se_h$
$Se_h = \{E_1, E_2, \dots, E_j\}$
トピック $T_k$ を条件とした文章 $Se_h$ の出現確率
$P(Se_h T_k) = \sum_j P(Se_h E_j)P(E_j T_k)$
係り受け $E_j$ が出現する中で文章 $Se_h$ が出現する確率( $E_j$ の出現文章数の逆数)
$P(Se_h E_j) = 1/n(E_j)$
トピック $T_k$ を条件とした文章 $S_h$ の出現確率 ※ $P(S_h Sw_h)$ と $P(S_h Se_h)$ はともに1/2とする
$P(S_h T_k) = P(S_h Sw_h)P(Sw_h T_k) + P(S_h Se_h)P(Se_h T_k)$
文章 $S_h$ の出現確率
$P(S_h) = \sum_k P(S_h T_k)P(T_k)$

## トピックスコア算出プロセス

### ①文章ごとにスコアを計算

特許ID	文章ID	T01	T02	T03	...	T47
1	1	3.1	0.9	2.0		1.1
1	2	1.4	0.2	5.5		2.4
2	1	0.8	5.8	1.3		0.9
2	2	1.2	3.2	1.7		1.0
2	3	0.6	1.8	2.6		3.6
...						

### ②特許IDごとに文章スコアを集約

※最大値を採用する

特許ID	T01	T02	T03	...	T47
1	3.1	0.9	5.5		2.4
2	1.2	5.8	2.6		3.6
...					

### ③閾値を設定してフラグに変換する

※閾値は3に設定する

特許ID	T01	T02	T03	...	T47
1	1	0	1		0
2	0	1	0		1
...					

# トピックのフラグデータの作成

全特許データに対して各トピックのスコア(該当有無)を計算することで、トピックをベースとした様々な分析を実行することができます

トピックのスコア(フラグ情報)を紐づけた特許データ

特許ID	出願番号	出願年	出願人	要約文		用途トピック U01	用途トピック U02	...	用途トピック U25	技術トピック T01	技術トピック T02	...	技術トピック T47
				【課題】	【解決手段】								
1	特願2006-XXXX	2006	A社	空気調和機の高外気	吸気口から導入された	1	0		0	0	1		0
2	特願2009-XXXX	2009	B社	短時間で除霜を行うこ	着霜検出手段が室外	0	1		0	1	0		0
3	特願2011-XXXX	2011	C社	乾燥運転が中断され	通風路を通して回転	0	0		1	1	0		0
4	特願2013-XXXX	2013	D社	ウインドシールドの防	車両用空調装置の空	0	1		0	0	1		1
...	...	...	...			...	...		...	...			...
30039	特願2012-XXXX	2012	Z社	プリ空調時に、除菌ま	冷暖房空調ユニットは	0	1		0	1	1		0

①出願年の分析

②出願人の分析

③用途と技術の関連性の分析

トピックをベースにした分析によって読むべき特許文書を効率的に絞り込むことができる

## 2. Nomolyticsを適用した特許分析事例

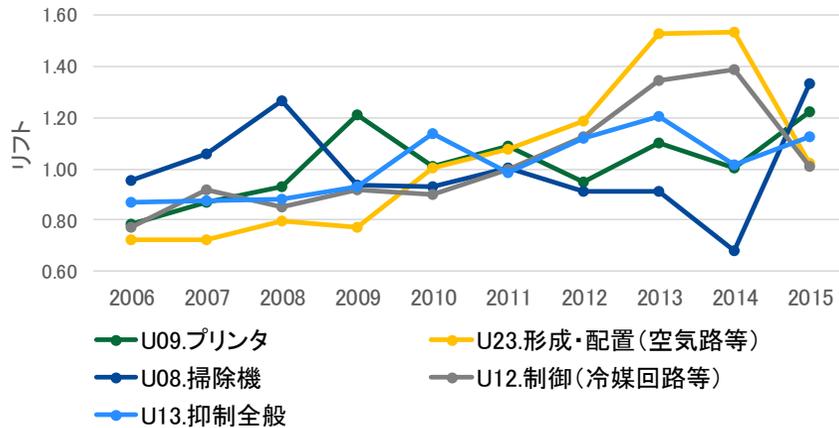
### 2-4. 出願年×トピックによるトレンド分析

#### 【分析目的】

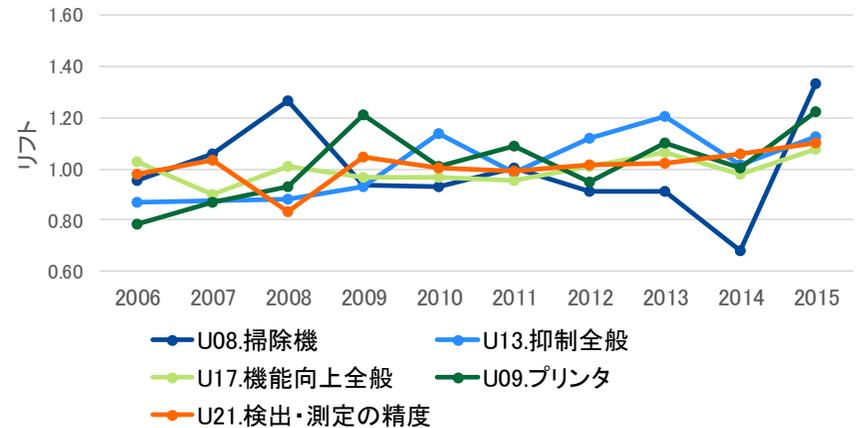
技術や用途のトレンドを把握し、有望なシーズやニーズを探り、今後の技術開発戦略を検討する

近年は掃除機や空気浄化、塵埃除去、プリンタに関する用途が上昇しています

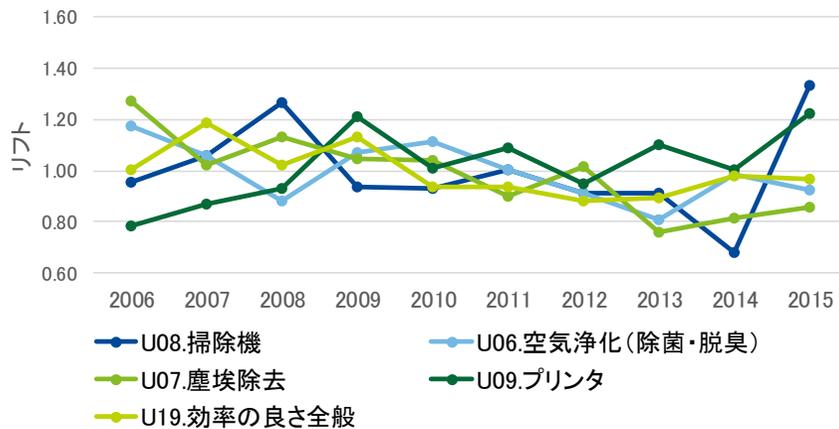
## 2006年からの上昇率 best5



## 2011年からの上昇率 best5



## 2013年からの上昇率 best5



## 集計の仕方

- リフト値を出願年・トピックごとに集計

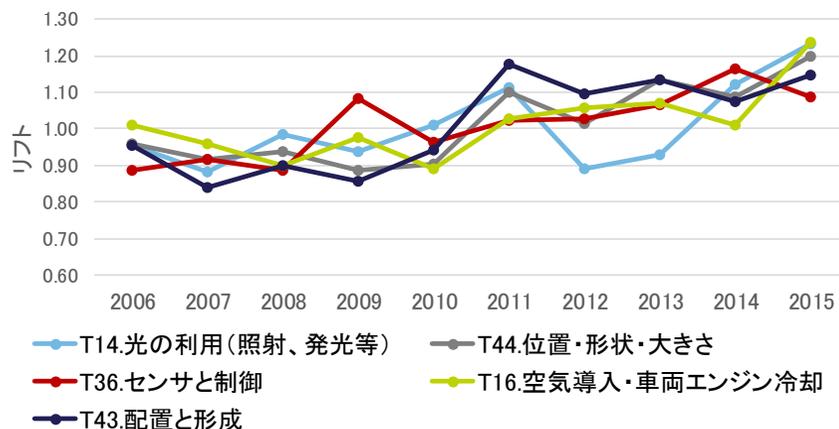
$$P(\text{出願年} | \text{トピック } U_x = 1)$$

$$P(\text{出願年})$$

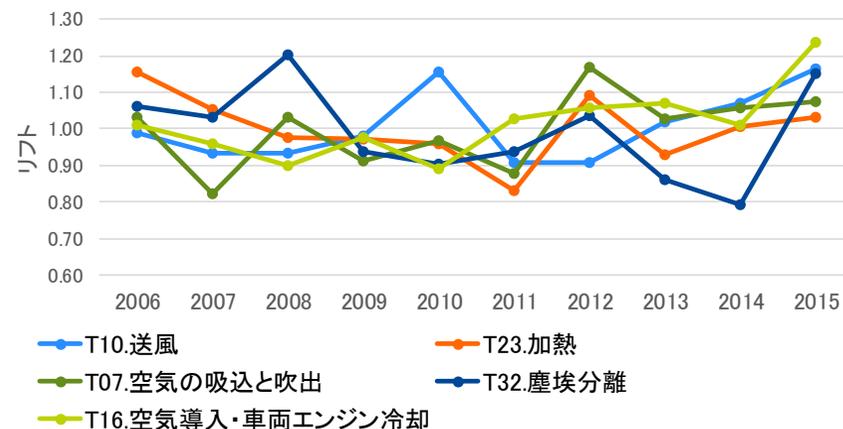
- その出願年の出願件数割合を平均(=1)として標準化した値

近年は塵埃分離や車両エンジンの冷却に関する技術が、長期的にはプロジェクトなどの光の利用に関する技術が上昇しています

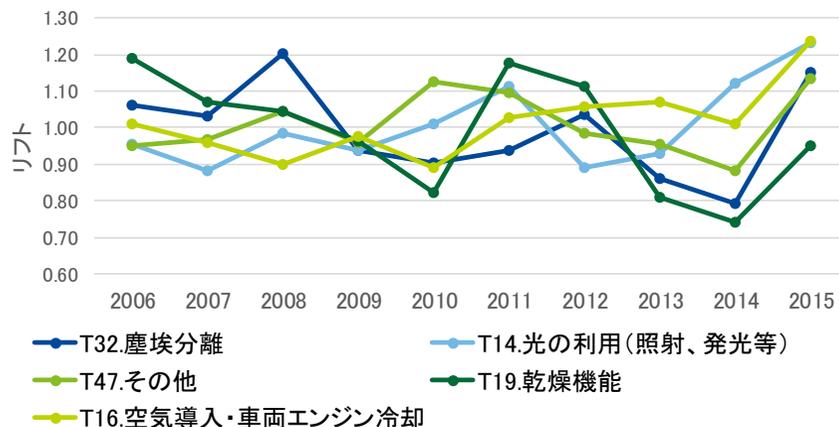
## 2006年からの上昇率 best5



## 2011年からの上昇率 best5



## 2013年からの上昇率 best5



## 集計の仕方

- リフト値を出願年・トピックごとに集計

$$P(\text{出願年} | \text{トピック } T_x = 1)$$

$$P(\text{出願年})$$

- その出願年の出願件数割合を平均(=1)として標準化した値

## 2. Nomolyticsを適用した特許分析事例

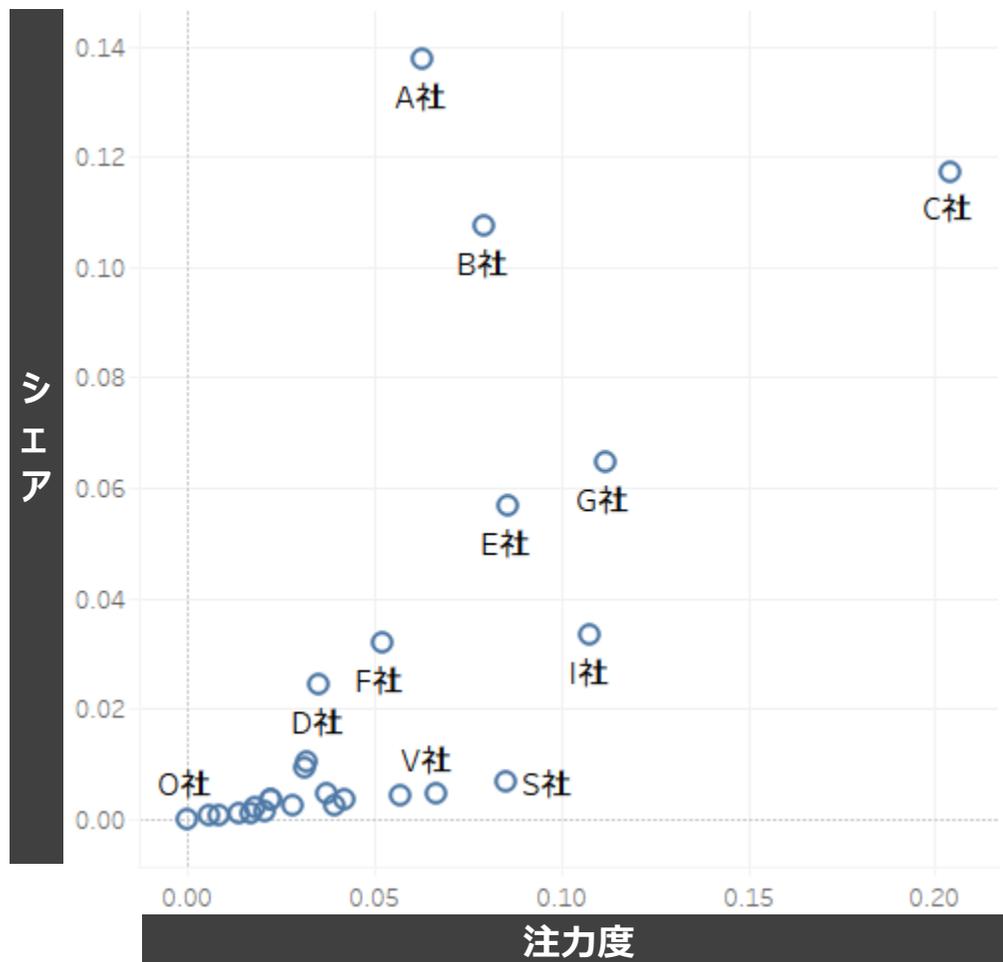
### 2-5. 出願人×トピックによる競合分析

#### 【分析目的】

各出願人の動向や、業界における棲み分け、競合他社と自社との関係性などを把握し、他社との差別化戦略や協業戦略を検討する

塵埃分離に関する技術は、1社の注力度が高いものの、他にもある程度のシェア・注力度を保有する企業が何社か存在するため、今後連携などの動きも考えられる領域と思われます

## 注力度とシェアの散布図



## 考察と戦略の検討

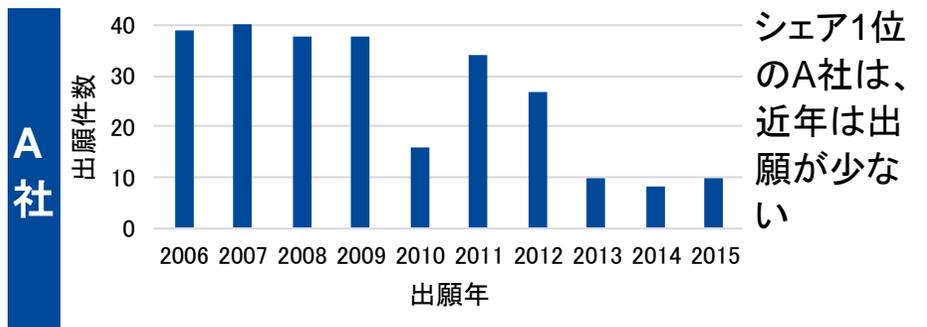
- C社は、高めのシェアを獲得しつつ、他社と比べて注力度がとても高く、高い技術力を保有していると考えられ、今後はよりシェアを伸ばすことで高シェア高注力度のポジションを確立することができると思われる
- A社とB社は、シェアは高いがまだC社に注力度で劣っているので、例えば規模は中程度だが注力度は比較的高く、技術力があると思われるE社、G社、I社などと連携することで、C社の上のポジションを狙うことができる可能性がある

## 注力度とシェア

- **注力度:**  $P(\text{トピック}T \mid \text{出願人}X)$ 
  - 出願人Xの出願特許の中で、どれくらいの割合がそのトピックTに該当するものか、つまり出願人がどれくらいそのトピックに注力しているのかを示している
- **シェア:**  $P(\text{出願人}X \mid \text{トピック}T)$ 
  - トピックTが該当する特許の中で、どれくらいの割合がその出願人Xの出願によるものか、つまりトピックのなかでどれくらいその出願人が占めているのかを示している

注力度1位のC社は直近で出願が急増し、高シェアのA社とB社の近年の出願動向は、A社は減少ですがB社は増加で、またシェア4位のG社も出願を伸ばしており、今後に要注目です

## 注目企業の出願件数のトレンド



## 注目企業の出願件数のトレンド



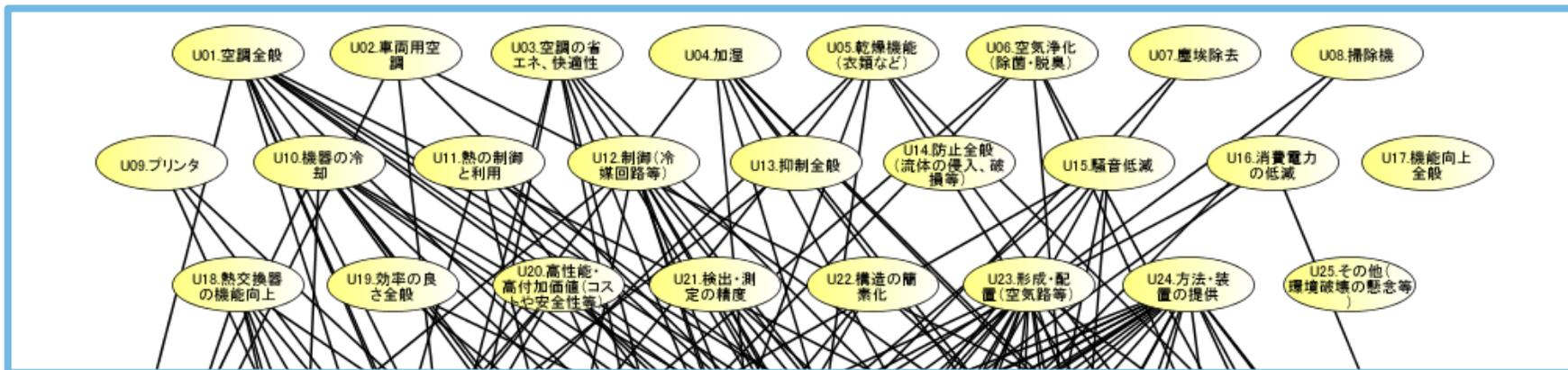
## 2. Nomolyticsを適用した特許分析事例

### 2-6. 用途×技術の関係分析<その1> ～用途⇒技術の関係～

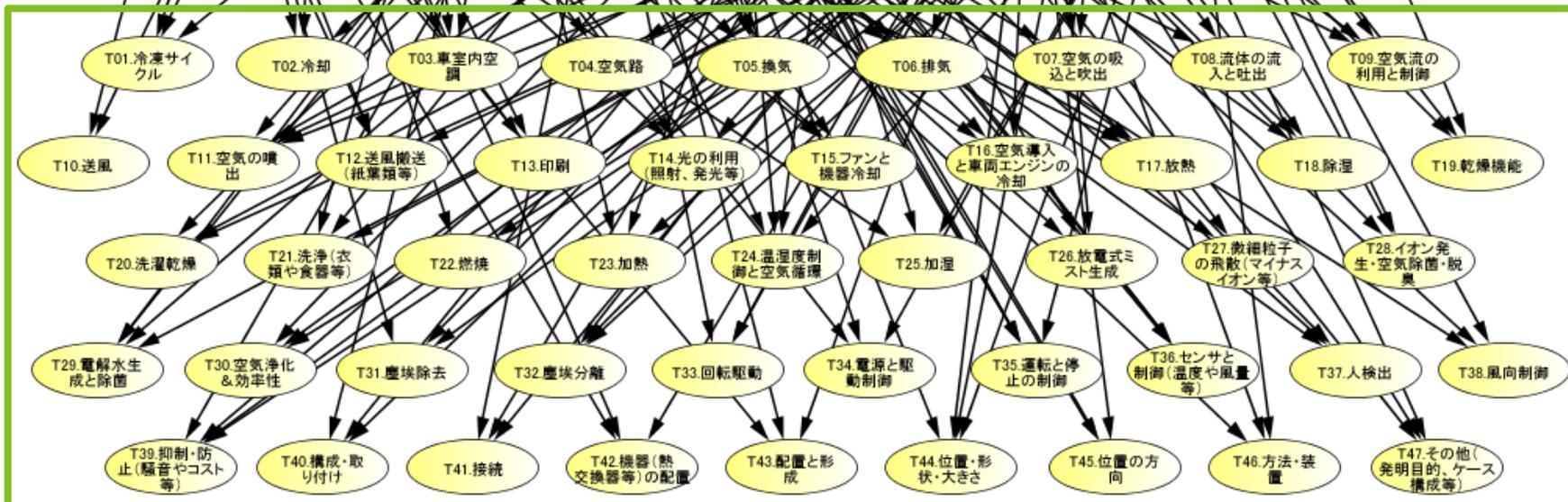
#### 【分析目的】

自社で検討中の事業実現のために重要な解決技術や代替技術、競合他社の存在を把握し、事業化のための開発戦略や他社との協業戦略を検討する

## ベイジアンネットワークを適用して、用途トピックに対する技術トピックの確率的因果関係をモデル化します



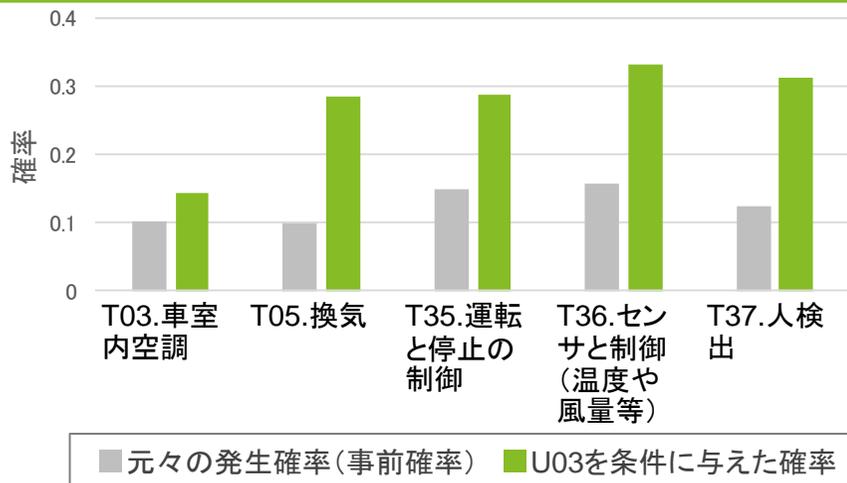
用途トピック



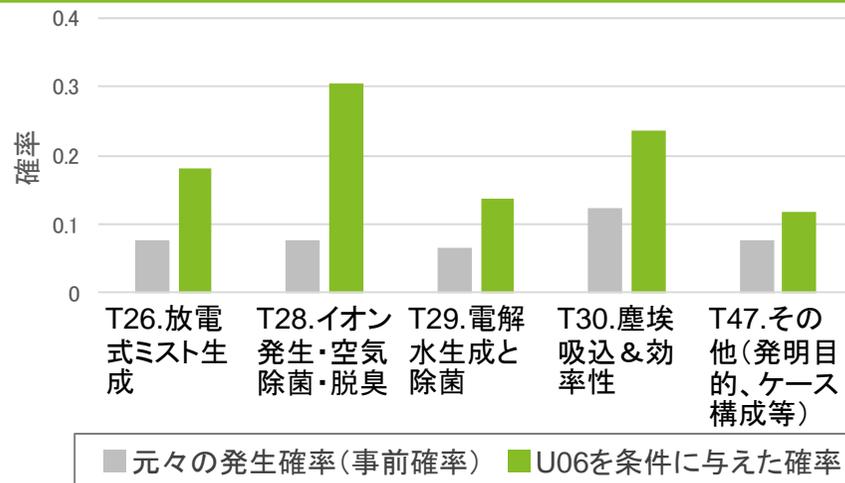
技術トピック

1つの用途トピックを条件に与えたときに、ベイジアンネットワークで見られた各技術トピックの確率の変化をシミュレーションし、用途に対する技術の関連性の強さを確認します

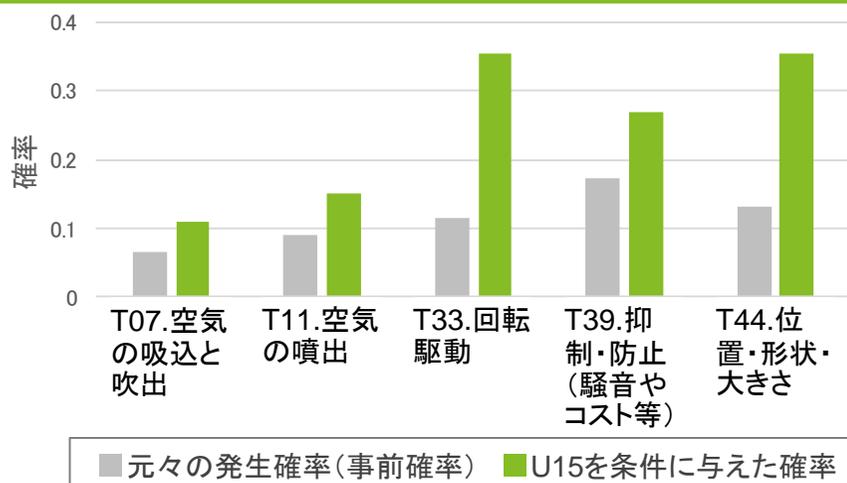
## 「U03.空調の省エネ、快適性」と関係のある技術



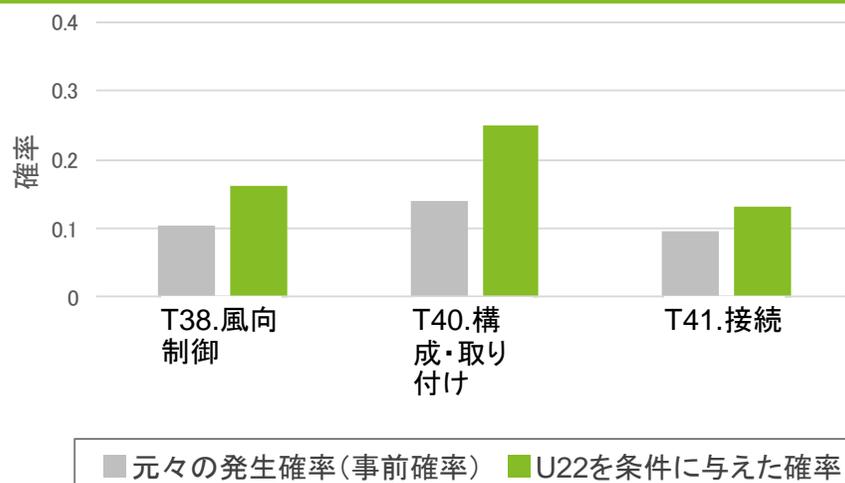
## 「U06.空気浄化(除菌・脱臭)」と関係のある技術



## 「U15.騒音低減」と関係のある技術

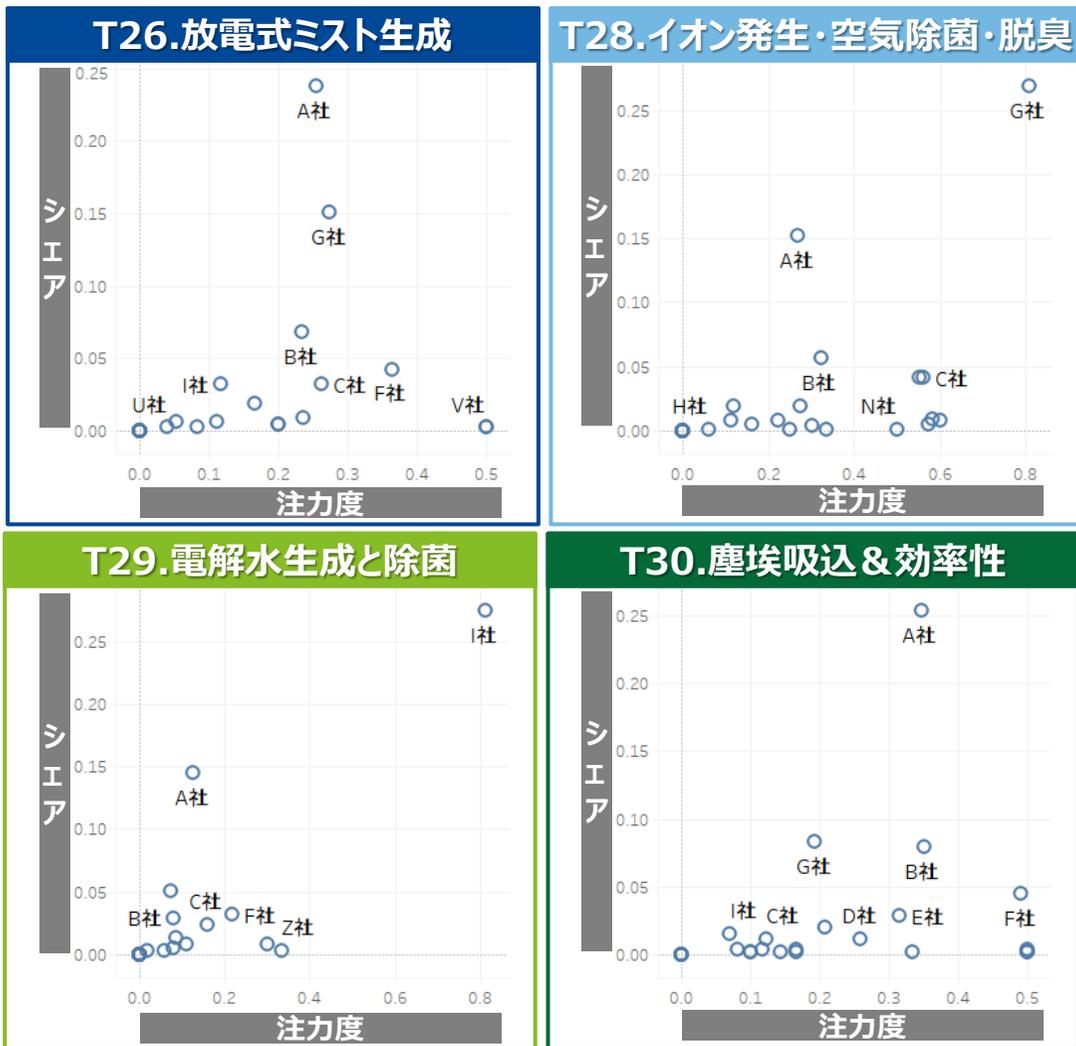


## 「U22.構造の簡素化」と関係のある技術



U06の用途と関係する4つの技術のうち2つは一強状態にあり、U06の事業化では、この技術を避けた他の技術の開発を検討する、あるいはその一強企業の買収も考えられます

## 「U06.空気浄化」の関係技術トピックにおける出願人マップ



## 考察と戦略の検討

- 「T28.イオン発生・空気除菌・脱臭と「T29.電解水生成と除菌」は、それぞれG社とI社が高シェア高注力度のポジションを確立した一強状態の技術といえる
- 「T26.放電式ミスト生成」と「T30.塵埃吸込&効率性」は、高シェア高注力度のポジションは空いているが、どちらもF社がシェアは低いものの注力度が高いポジションにある
- 一強状態の技術を避けて「U06.空気浄化」の用途を実現する場合、T26やT30の技術が狙い目といえるが、注力度の高いF社は要注目である
- 一強状態にあるT28やT29の技術において、その一強企業と提携あるいはM&Aを実現すれば、その技術領域ごと獲得できる

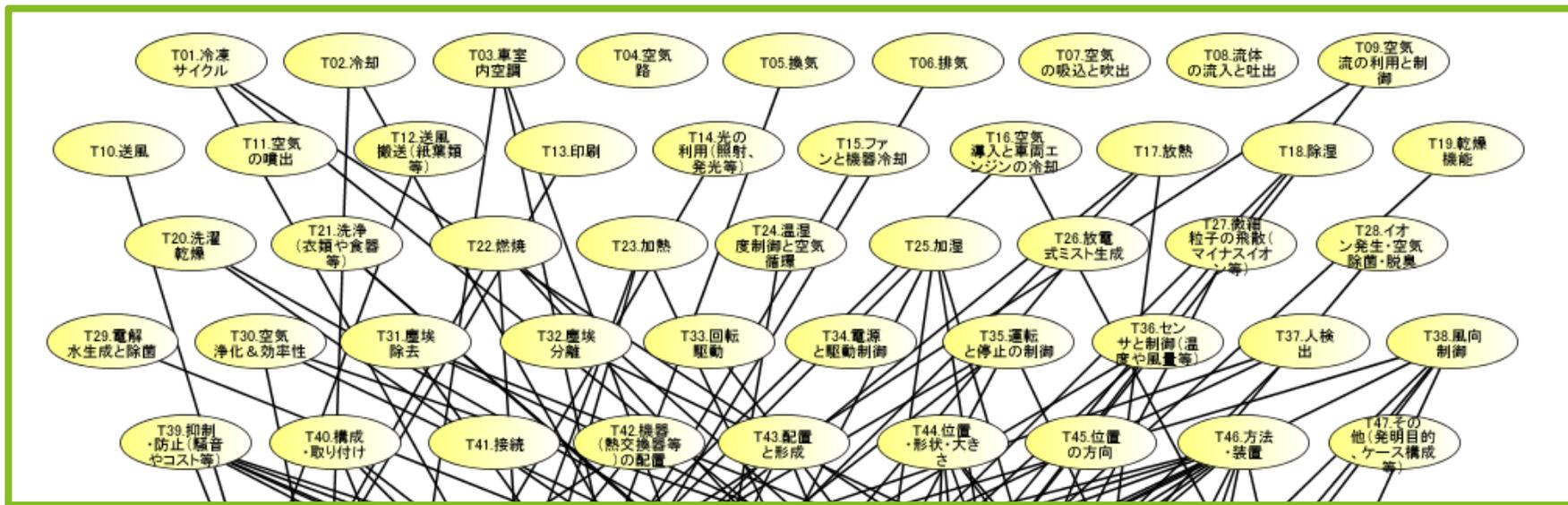
## 2. Nomolyticsを適用した特許分析事例

### 2-7. 用途×技術の関係分析<その2> ～技術⇒用途の関係～

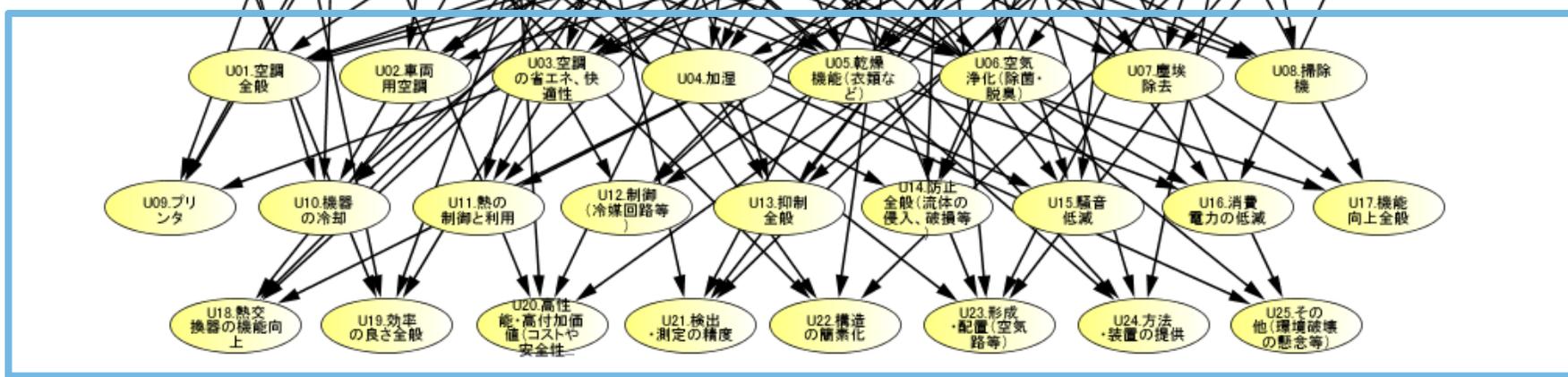
#### 【分析目的】

自社技術と関係のある用途を把握し、まだ自社で想定していない用途を見つけ、保有技術を有効活用できる新しい用途展開のアイデアを創出する

## ベイジアンネットワークを適用して、技術トピックに対する用途トピックの確率的因果関係をモデル化します



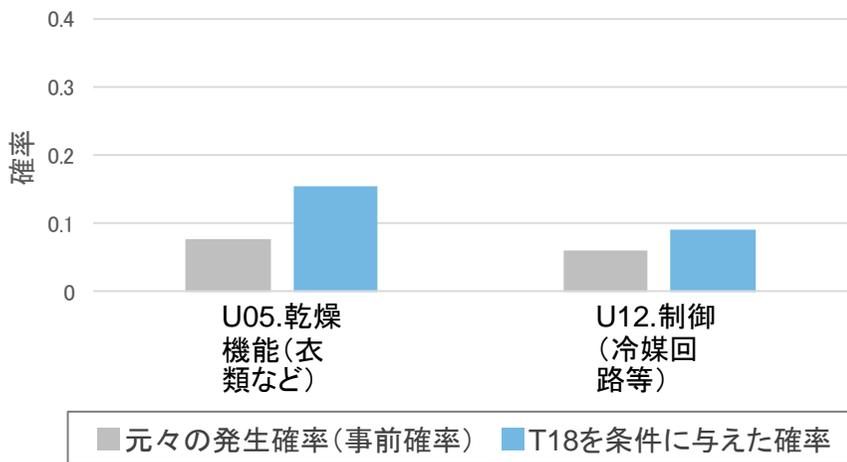
技術トピック



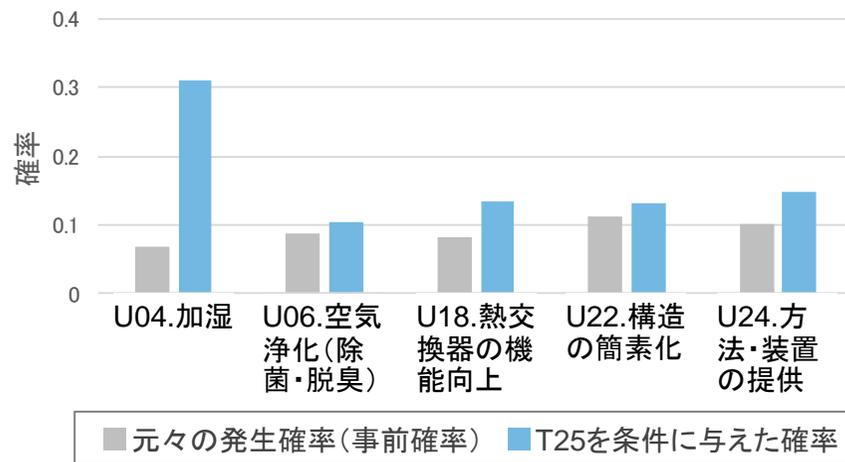
用途トピック

1つの技術トピックを条件に与えたときに、ベイジアンネットワークで関係が見られた各用途トピックの確率の変化をシミュレーションし、技術に対する用途の関連性の強さを確認します

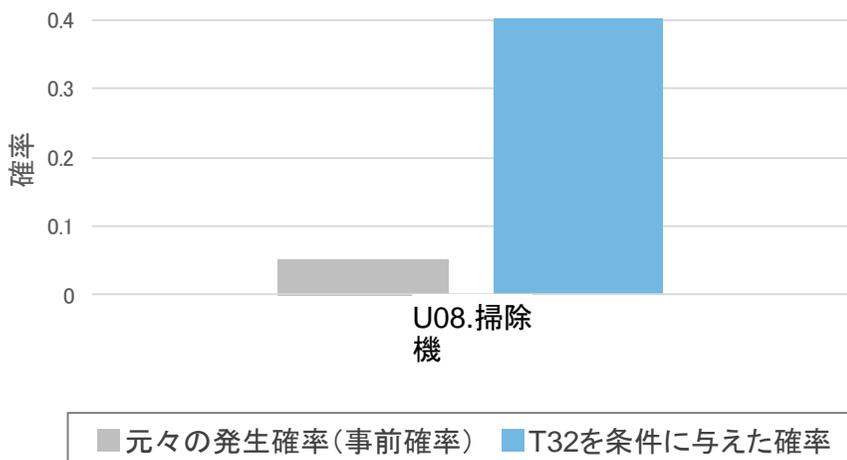
## 「T18.除湿」と関係のある用途



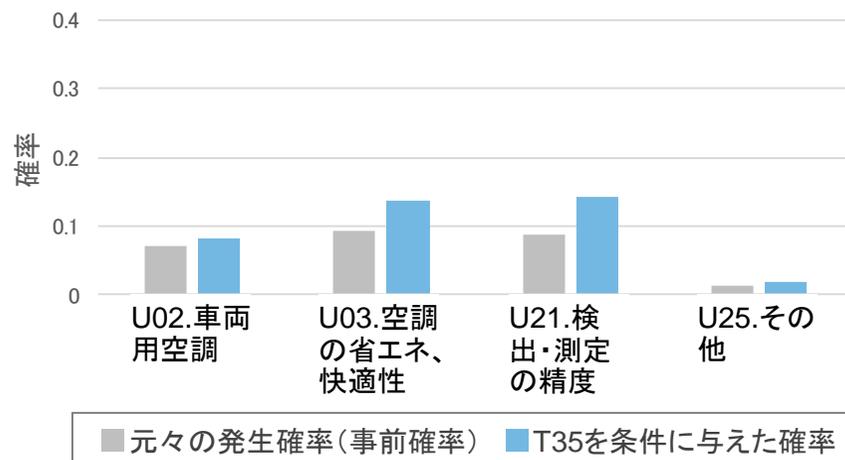
## 「T25.加湿」と関係のある用途



## 「T32.塵埃分離」と関係のある用途

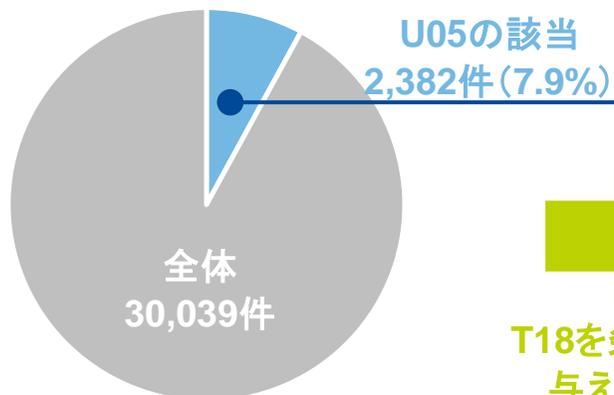


## 「T35.運転と停止の制御」と関係のある用途

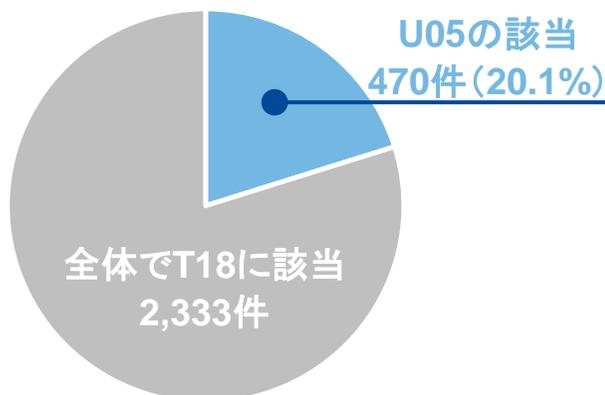


「T18.除湿」の技術の応用先として「U05.乾燥機能」の用途は高い関連性がありますが、出願人Wの保有するT18ではそれがほとんどなく、新規用途となる可能性があります

## 全体でのU05の該当割合



## 全体でのT18におけるU05の割合



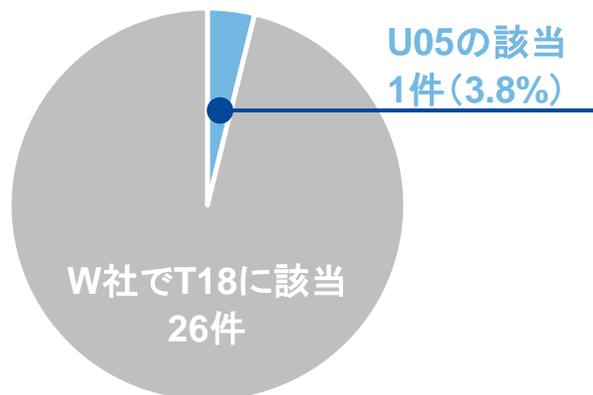
T18を条件に  
与えると

## T18 × U05の用途語

T18 × U05に該当する特許の【課題】文章で特徴的に出現する用途ワード

乾燥性能 デシカントロータ 乾燥工程  
乾燥運転 除湿ロータ 浴室乾燥機  
短縮 乾燥装置 乾燥用空気 被乾燥物  
乾燥+できる 効率 洗濯乾燥機  
衣類除湿機 除湿装置  
衣類乾燥機時間 乾燥 洗濯物  
除湿能力 乾燥機 除湿  
ヒートポンプ 乾燥時間 ドラム式洗濯乾燥機 乾燥手段  
乾燥効率 乾燥度合い

## 出願人WのT18におけるU05の割合



## 考察

- ベイジアンネットワークのモデルでは、技術「T18.除湿」に対する用途「U05.乾燥機能(衣類など)」の関係が見られた
- 全体では、U05の該当は7.9%だが、T18を条件としたときでは、その該当割合が20.1%となり高い関連性が認められる
- しかし、出願人Wでは、T18に該当する特許のうち、U05に該当する特許は1件だけである  
⇒W社の保有するT18の技術はU05の用途への展開も考えられる

## W社の印刷機の中でインク液を吸収した用紙の湿気をムラなく取り除く除湿技術は、洗濯乾燥機の中で洗濯物をムラなく効率的に乾燥させることにも応用できるかもしれません

### T18がU05で応用されている例

#### 発明の名称

ドラム式洗濯乾燥機

#### 【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。

### 出願人Wの保有するT18の例

#### 発明の名称

インクジェット記録装置及び画像記録方法

#### 【課題】

処理液の厚みムラを低減するとともに処理液による用紙のコックリングを低減することで、高品質かつ高速の画像記録を可能とするインクジェット記録装置及び画像記録方法を提供する。

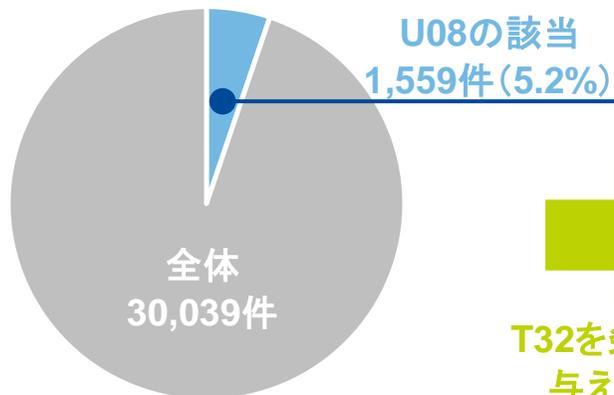
#### 【解決手段】

記録媒体に処理液を付与する処理液付与部の後段には、記録媒体表面に残存する溶媒を蒸発させるプレ加熱部が設けられている。プレ加熱部はIRプレヒータにより記録媒体表面を輻射加熱するとともに、吸引ファンにより記録媒体表面の湿り空気を置換する。液状の処理液が不均一にならないように乾燥処理を施すことで、均一な膜厚を持つ固体状の凝集処理層が形成される。その後、本加熱部による熱風噴射加熱により、コックリング量が所定量以下になるように本加熱処理が施される。

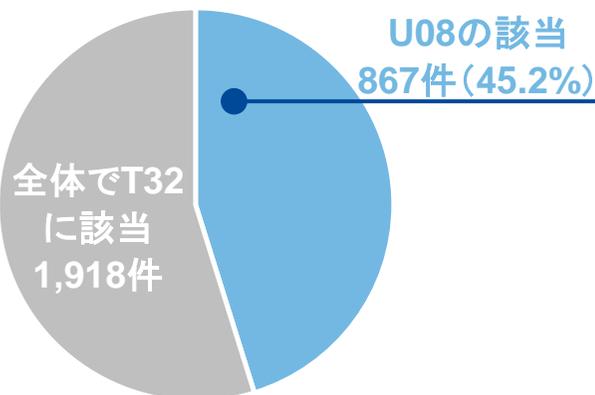
※対外説明用のため要約文は一部加工している

「T32.塵埃分離」の技術の応用先として「U08.掃除機」の用途はとても高い関連性がありますが、出願人Sの保有するT32ではそれが全くなく、新規用途となる可能性があります

全体でのU08の該当割合



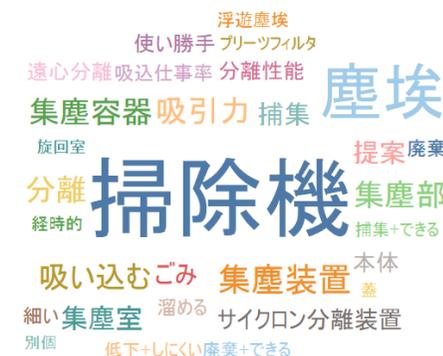
全体でのT32におけるU08の割合



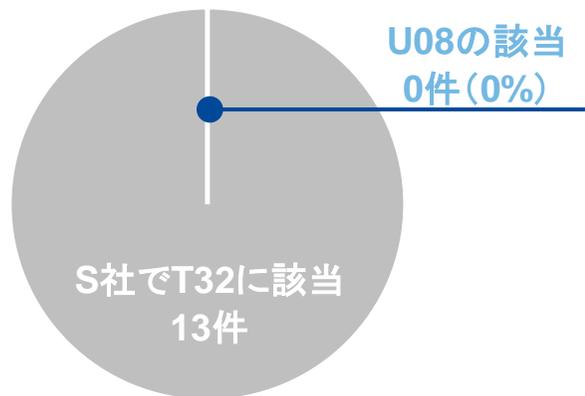
T32を条件に  
与えると

T32 × U08の用途語

T32 × U08に該当する特許の【課題】文章で特徴的に出現する用途ワード



出願人SのT32におけるU08の割合



考察

- ベイジアンネットワークのモデルでは、技術「T32.塵埃分離」に対する用途「U08.掃除機」の関係が見られた
- 全体では、U08の該当は5.2%だが、T32を条件としたときでは、その該当割合が45.2%となりとても高い関連性が認められる
- しかし、出願人Sでは、T32に該当する特許のうち、U08に該当する特許は0件である  
⇒S社の保有するT32の技術はU08の用途への展開も考えられる

## S社の印刷機でトナーを分離・回収するサイクロン部の清掃時期を判断して分離効率を維持する技術は、サイクロン掃除機の集塵部の集塵性能の向上にも応用できるかもしれません

### T32がU08で応用されている例

#### 発明の名称

電気掃除機

#### 【課題】

集塵性能が向上しメンテナンスの軽減が図れる電気掃除機を提供すること。

#### 【解決手段】

塵埃を含む空気を旋回させ塵埃分離する略円筒状の1次旋回室と、1次旋回室に連通した2次旋回室と、1次旋回室の下方に位置し塵埃を溜める集塵室と、塵埃を圧縮する圧縮板と、塵埃が流入する流入口を有し、圧縮板の底面の一部に突出部を流入口から見て集塵室の奥側に配設する構成としたことより、集塵室内に入った塵埃は、圧縮板の突出部に引っかかり動きが止められ、流れに乗って2次旋回室や1次旋回室側に戻ることが無いいため集塵性能が向上し、排気筒の詰まり防止によるメンテナンスの軽減を図ることができる。

### 出願人Sの保有するT32の例

#### 発明の名称

画像形成装置

#### 【課題】

サイクロン部の清掃時期を適正に判断して、トナーの分離効率の低下を抑制することが可能な画像形成装置を提供する。

#### 【解決手段】

画像形成装置は、トナー含有空気からトナーを遠心分離するサイクロン部と、サイクロン部によって分離されたトナーを回収する回収部と、サイクロン部によってトナーが分離された空気を通過させ、残留トナーを捕集するフィルタ部と、空気を吸引する送風部と、フィルタの汚れを検知する汚れ検知センサが設けられたトナー捕集部を備え、汚れ検知センサで検知されたフィルタの汚れから推定した風量と、風速センサで取得した風量の実測値の差分が、サイクロン清掃閾値を超えたと判断すると、サイクロン部の清掃モードを実行する。

※対外説明用のため要約文は一部加工している

### 3. Nomolyticsによる特許文書分析のまとめ

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

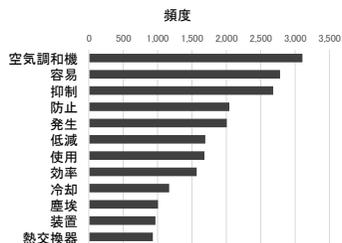
# Nomolytics®: Narrative Orchestration Modeling Analytics

## 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

単語抽出

Text Mining Studio



## 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

トピック抽出

Visual Mining Studio

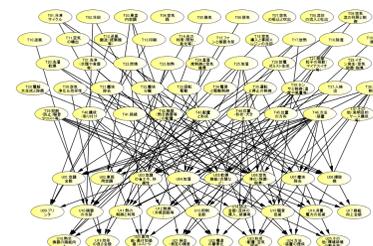


## ベイジアンネットワーク

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

モデリング

BayoLink



膨大なテキストデータを人間が理解しやすい形に整理できる

テキスト情報内に潜む複雑な要因関係を構造化できる

条件を変化させたときの結果の挙動をシミュレーションできる

ある事象の発生確率をコントロールする条件を発見できる





用途と技術の統計的な関係を把握することで、各用途を実現するための重要技術を確認して技術戦略を検討したり、自社技術を有効活用できる新規用途のアイデアを創出できます

## 従来の特許分析

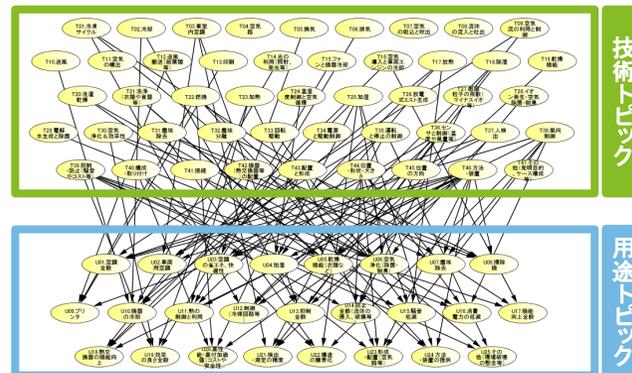
課題と解決手段のカテゴリ間のクロス集計

課題	解決手段カテゴリ									
	カテゴリ01	カテゴリ02	カテゴリ03	カテゴリ04	カテゴリ05	カテゴリ06	カテゴリ07	カテゴリ08	カテゴリ09	カテゴリ10
カテゴリ01	206	80	71	184	26	47	11	9	43	1
カテゴリ02	208	76	87	182	23	48	9	15	40	2
カテゴリ03	172	74	53	57	31	35	10	21	20	3
カテゴリ04	176	54	37	59	26	46	29	26	9	5
カテゴリ05	85	39	13	23	14	16	5	0	7	2
カテゴリ06	87	53	31	33	59	37	15	24	28	19
カテゴリ07	79	68	82	28	24	12	6	16	18	15
カテゴリ08	32	29	19	1	20	5	17	2	4	2

- 【課題】と【解決手段】それぞれに対して人がグルーピングして作成したカテゴリのクロス集計表を作成し、その対応関係を考察する
- その組み合わせで出願件数が多いからといって、統計的に意味のある関係であるとは限らない(全体的に出願件数が多い可能性もある)

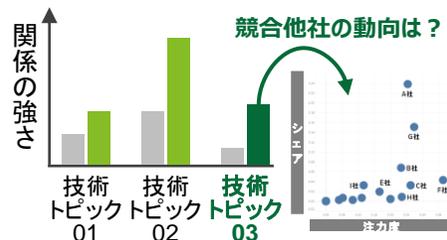
## Nomolyticsを適用した特許分析

課題と解決手段のトピック間の統計的な因果関係モデル

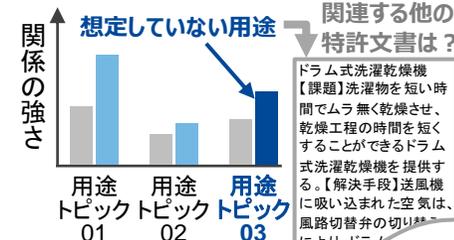


分析目的に応じて因果構造を入れ替える  
 ①「技術 ⇒ 用途」  
 ②「用途 ⇒ 技術」

想定用途と関係のある技術



自社技術と関係のある用途



- 客観的に抽出されたトピックをベースに課題と解決手段(用途と技術)の統計的な関係をベイジアンネットワークで把握できる
- 検討中の用途に対して、関係の強い技術を確認し、各技術における出願人の動向から自社の技術戦略を検討できる
- 自社技術と関係の強い用途で想定していないものを確認し、その関連特許の探索から技術の新規用途アイデアを創出できる

## 4. 新たなトピックモデル技術:PCSA

PLSAでは、テキストマイニングで全体のテキストデータから抽出された単語で構成される共起行列をインプットにすることで、データ全体を表現する平均的なトピックを抽出します



テキストデータにテキストマイニングを実行し、全体のデータに含まれる単語を抽出する

抽出した各単語の全体における共起頻度を集計した共起行列を作成する

作成した共起行列にPLSAを適用し、使われ方の似ている単語をトピックに集約する

ターゲット事象の該当データと非該当データからそれぞれ構築した2つの共起行列の差分にPLSAを適用することで、そのターゲット事象に影響を与えるトピックを優先して抽出します

# PCSA<sup>®</sup> (Probabilistic Causal Semantic Analysis: 確率的因果意味解析)



単語抽出

共起頻度集計

トピック抽出



出現頻度

- Word 1
- Word 2
- Word 3
- Word 4
- Word 5
- Word 6
- Word 7
- Word 8
- Word 9
- Word x

あるターゲット事象(属性情報)Xが、

全データ

共起行列U

該当するデータ

該当しないデータ

共起行列A

共起行列B

2つの共起行列の差分

共起行列A-B

PLSAの適用

トピック1

Word5  
Word3 Word9  
**Word1**  
Word2 Word8  
Word4 Word6  
Word7

トピック2

Word2 Word8  
**Word3**  
Word1 Word9  
Word7 **Word6**  
Word5 Word4

トピック3

Word6  
**Word2**  
Word3 **Word5**  
Word8 Word1 Word9  
**Word7**  
Word4

ターゲット事象Xの該当有無に影響を与える要因トピックを優先的にテキスト情報から抽出できる

ターゲット事象 X

全データから構築した共起行列Uを、あるターゲット事象(属性情報)Xが該当するデータから構築した共起行列Aと、該当しないデータから構築した共起行列Bに分割し、その2つの共起行列の差分を取った共起行列(A-B)に対してPLSAを適用する

人工知能学会  
2018年度全国大会  
優秀賞 受賞

Nomolyticsは全体を表すトピックを抽出してから属性との関係进行分析しましたが、PCSAはその属性の特徴の要因となるトピックを最初から抽出し、より効果的な特徴の探索を行います

## Nomolyticsでは

データ全体を表現するトピックを抽出

してから

属性との関係进行分析

すると

属性の特徴を把握

できた

## PCSAでは

属性の特徴をよく表現する

ような

偏ったトピックを抽出

してから

より顕著な属性との関係进行分析

する

## 5. PCSAを適用した特許分析事例

## 5. PCSAを適用した特許分析事例

### 5-1. 「車・電気」に関する特許文書データ

「車」「電気」を含む10年分の特許データ26,419件の要約文を対象に、PLSAで全体のトピックを、PCSAで出願年・パテントスコア・出願人の属性で特徴を示すトピックを抽出します

## 分析データの抽出条件

- 対象: 公開特許公報
- キーワード: 要約と請求項に「車」と「電気」を含む
- 出願日: 2007年1月1日～2016年12月31日
- 抽出方法: Patent Integrationを使用
- 抽出件数: 26,419件



## 分析対象

- トピック抽出対象: 要約文のテキスト情報
- 使用する属性情報
  - 出願年: 2007年～2016年
  - PISコア  
請求項の広さ、被引用回数に比例する指標  
(Patent Integrationの独自評価指標)
  - 出願人: 出願件数上位26社 (個人は除く)  
※グループ会社を統一するなどの名寄せ済み

## トピックの抽出対象

### ① PLSAで抽出する全体の集約トピック



▶ データ全体を表すトピックの分類を把握する

### ② PCSAで抽出する属性別の特徴トピック



▶ 最近の出願特許の特徴を把握する



▶ スコアの高い有用技術の特徴を把握する



▶ 出願数の多い企業とそうでない企業の棲み分けを把握する

## 5. PCSAを適用した特許分析事例

### 5-2. PLSAによる全体のトピックの抽出

該当件数: 26,419件

## テキストマイニングで単語と係り受け表現を抽出し、単語×係り受けで構成される共起行列にPLSAを適用することで単語と係り受けの出現の背後にある潜在トピックを抽出します

### テキストマイニングの実行

要約文に含まれる「単語(名詞)」と「係り受け」を抽出する

単語	頻度
構成	4,997
制御	4,360
配置	3,895
モータ	3,486
形成	3,459
供給	3,309
検出	3,215
電気自動車	3,181
...	...

係り受け表現	頻度
電力⇒供給	1,208
否⇒判定	517
モータ⇒駆動	460
バッテリー⇒充電	440
効率⇒良い	419
供給⇒電力	332
電気自動車⇒提供	285
充電⇒行う	273
...	...

### 共起行列の作成

「単語×係り受け」の共起行列(文章単位で同時に出現する頻度のクロス集計表)を作成する

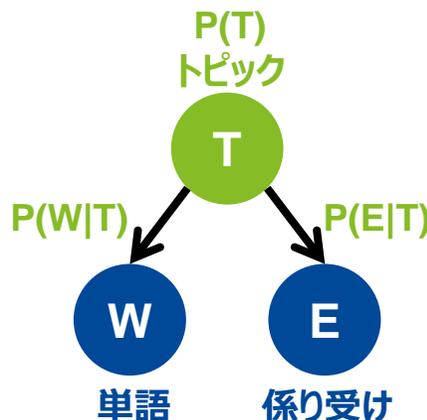
	係り受け表現				
	電力↓供給	否↓判定	モータ↓駆動	バッテリー↓充電	...
構成	118	33	36	33	
制御	268	73	108	85	
配置	69	2	29	8	
モータ	239	61	494	58	
...					

単語

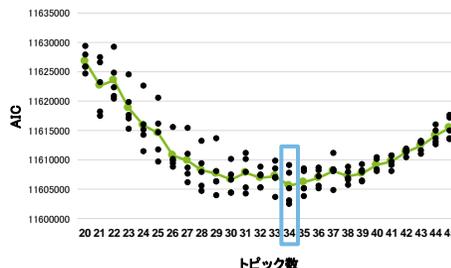
単語(名詞): 3,020語  
係り受け: 2,128表現  
※頻度20件以上を対象

### PLSAの実行

共起行列にPLSAを適用する



トピック数を幅を持たせて設定し、各トピック数に対してPLSAを初期値を変えて5回ずつ実行して情報量基準AICを計算し、AIC最小の解を採用する



### トピックの抽出

各トピックについて以下の3つの確率が計算される

- ①  $P(T)$   
トピックの存在確率
- ②  $P(W|T)$   
トピックにおける単語の所属確率
- ③  $P(E|T)$   
トピックにおける係り受けの所属確率

トピックにおける $P(W|T)$ と $P(E|T)$ からトピックの意味を解釈する

トピック Tu13	
P(T) = 5.0%	

P(W T)	単語	P(E T)	係り受け
12.6%	充電	5.1%	バッテリー充電
8.9%	電気自動車	4.0%	充電-行う
6.5%	蓄電装置	3.9%	電気自動車-充電
3.0%	バッテリー	1.9%	蓄電池-充電
2.0%	充電システム	1.6%	蓄電装置-充電
2.0%	蓄電池	1.6%	電力-供給
1.9%	電力	1.3%	充電-開始
1.7%	制御	1.2%	電気自動車-接続
1.5%	充電スタンド	1.2%	充電-蓄電装置
1.5%	放電	1.1%	用いる-充電
...	...	...	...

確率の高い構成要素から、Tu13のトピックは「電気自動車の蓄電池充電」に関するトピックと解釈できる



# 全体トピック34個の一覧②

26,419件の特許は、エンジン、動力伝達、モータ、ブレーキ、電力変換、二次電池、充電、情報通信、異常検出、筐体、構成、小型化、安全性などの34個のトピックに集約されました

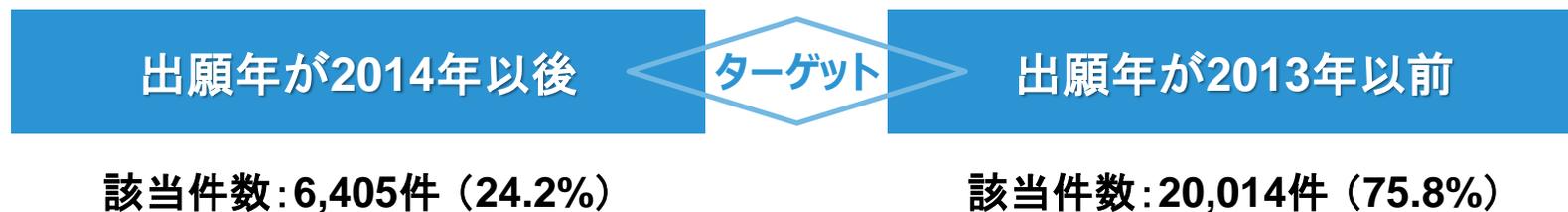


Tu32, Tu33, Tu34は 特定の出願人による 重複した要約内容の 特許から抽出された

※文字の大きさはトピックに対する関係の強さを表現している（上位5つの単語を赤色で表示している）

## 5. PCSAを適用した特許分析事例

### 5-3. PCSAによる出願年の特徴トピックの抽出



分析の  
ポイント

最近ホットなトピックとは？  
もうトレンドが終わっているトピックとは？

## ① 2013年以前データ、② 2014年以後データで作成した2つの共起行列の差分にPLSAを適用することで、2014年前後の出願に特徴がある技術トピックを優先的に抽出します

### 2つの共起行列の作成

①出願年が**2013年以前**のデータ  
20,014件(文章数:33,113件)

#### 係り受け表現

	電力↓供給	否↓判定	モータ↓駆動	バッテリー↓充電	...
構成	97	30	27	25	
制御	208	49	80	66	
配置	56	1	23	8	
モータ	192	42	356	46	
...					

単語

②出願年が**2014年以後**のデータ  
6,405件(文章数:10,723件)

#### 係り受け表現

	電力↓供給	否↓判定	モータ↓駆動	バッテリー↓充電	...
構成	21	3	9	8	
制御	60	24	28	19	
配置	13	1	6	0	
モータ	47	19	138	12	
...					

単語

### 差分の共起行列の作成

①2013年以前の共起行列と  
②2014年以後の共起行列の  
差の絶対値を計算した共起行列  
を作成する

#### 係り受け表現

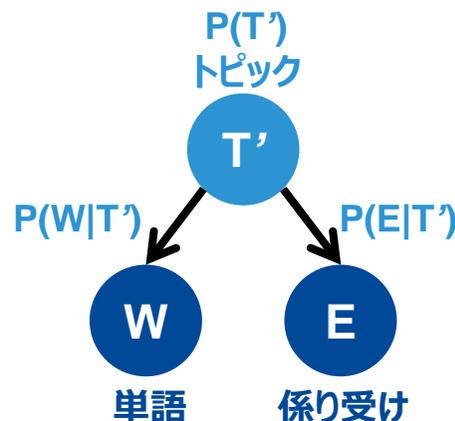
	電力↓供給	否↓判定	モータ↓駆動	バッテリー↓充電	...
構成	10.4	6.7	0.3	0.1	
制御	7.4	8.1	2.1	2.4	
配置	5.1	0.7	1.4	2.6	
モータ	15.2	5.4	22.7	2.9	
...					

単語

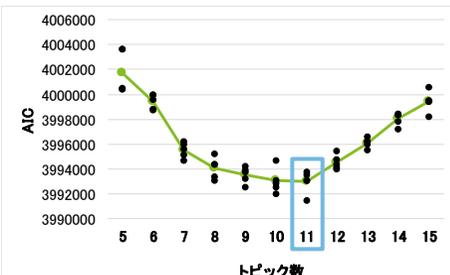
ただし、2つの共起行列は異なる  
文章数のデータから作成されて  
いるので、①2013年以前の共起  
行列の頻度を2つの文章数の比  
率(10,723/33,113)で重み調整  
してから、②2014年以後の共起  
行列との差を計算する

### PLSAの実行

差分の共起行列にPLSAを適用  
する



トピック数を幅を持たせて設定し、  
各トピック数に対してPLSAを初  
期値を変えて5回ずつ実行して  
情報量基準AICを計算し、AIC最  
小の解を採用する



### トピックの抽出

各トピックについて以下の3つの  
確率が計算される

- ①  $P(T')$   
トピックの存在確率
- ②  $P(W|T')$   
トピックにおける単語の所属確率
- ③  $P(E|T')$   
トピックにおける係り受けの所属確率

トピックにおける $P(W|T')$ と $P(E|T')$   
からトピックの意味を解釈する

#### トピック Ta05

$P(T) = 2.5\%$

$P(W T)$	単語	$P(E T)$	係り受け
3.1%	製造	4.5%	水素-製造
2.2%	二次電池	3.0%	発電-電気
1.5%	負極	2.6%	製造方法-提供
1.5%	正極	2.6%	製造-方法
1.5%	製造方法	1.6%	含む-リチウムイオン電池
1.2%	エネルギー	1.3%	成形品-提供
1.1%	リチウムイオン電池	1.3%	強度-有する
1.1%	水素	1.3%	表面-形成
1.1%	セパレータ	1.2%	含む-組成物
1.0%	電解液	1.2%	リレー-スイッチ
...	...	...	...

確率の高い構成要素から、Ta05の  
トピックは「二次電池の製造方法」  
に関するトピックと解釈できる

# 出願年の特徴トピック11個の一覧

出願年(2014年前後)で特徴を示すトピックは、エンジン駆動、モータ構成、電力変換、充電、二次電池の製造方法、冷却・加熱、検出判定、小型化・低コスト化など11個抽出されました

### Ta01.エンジン駆動・動力伝達の制御

動力 演算 発電機  
駆動輪伝達クラッチ出力軸開始構成 **モータ**  
トランスミッション制御手段駆動運転者  
**エンジン 制御装置 ハイブリッド車両**  
判定 **制御** ブレーキ モータジェネレータ トルク  
連結 駆動力バッテリー停止 車輪 内燃機関  
作動 検出

### Ta02.モータの回転構成

位置 伝達 方向形成  
中心固定ハウジング自動車 **回転**  
収容移動できる 電気機械 駆動部  
**モータ支持回転** できる **配置 回転軸**  
軸方向 ステータ連結駆動周囲 移動口  
**構成** 車輪 発電機 シャフト 結合

### Ta03.交流・直流の変換

放電直列直流スイッチ電力 並列 蓄電池  
負荷 **供給 変換** バッテリー検出 直流電力  
交流電力充電電気自動車コンバータ **インバータ**  
**制御** 制御部駆動 電力変換装置 蓄電装置  
コンデンサ 制御装置 **電圧** 電流オフ電源モータ

### Ta04.電気自動車への充電、給電装置

検出判定走行 構成 外部  
充電ケーブル **電力** 制御 ユーザ 演算 取得電源  
収容蓄電池 **電気自動車 蓄電装置** コネクタ 充電+できる  
充電システム **充電 バッテリー**  
充電スタンド外部電源 **給電開始供給情報** 送信  
設置

### Ta05.二次電池の製造方法

由来 負極活物質 電解液 成形品  
セパレータエネルギー **正極** 電子機器正極活物質  
積層エネルギーシステム  
**製造 負極 二次電池** リチウムイオン電池  
再生可能エネルギー電気部品 **製造方法** スイッチ  
方法 自動車部品電池特性 水素 形成 安全  
電極含有表面発電 活物質

### Ta06.空調などの冷却・加熱

制御 燃料 配置 冷媒 排出 空気  
変換電気ヒータ **エンジン 電気エネルギー 供給 発電機**  
モータ内燃機関 **駆動 加熱 構成** 冷却水 **発電**  
バッテリー 燃料電池 制御装置 温度 **電力**  
排気ガス冷却 車室内 作動 熱 循環

### Ta07.情報通信、検出判定システム

電流送信自動車 調整 比較 情報  
信号受信 **構成** センサ 取得 制御装置  
システム **制御 検出** 速度 方法 制御部  
演算 **判定** 生成 測定電圧 電気信号  
動作 記憶 入力 変化 モータ 位置

### Ta08.形成・配置

開口 他端外部バッテリー 方向 下方  
開口部ケース 位置端子一端上方 **収容**  
コネクタ **固定 形成** 面ハウジング **配置**  
電気部品 基板 端部 接触 突出 **構成**  
保持 支持 対向 筐体 挿入

### Ta09.小型化・低コスト化・簡素化・操作性向上

精度 低コスト 効率 バッテリー 安定 **操作** 起因 安価  
コスト自動車 小型化車室内 損傷 ワイヤハーネス  
操作+できる **電気自動車** 確保ハイブリッド車両  
振動 **構成** 信頼性 **構造** 電子機器 **スイッチ**  
部品点数必要+ない モータ 影響  
実現安全

### Ta10.重力発電の活用による地球温暖化防止

駆動 タービン 既存火力発電 発電量 安価  
重力発電運用 燃料費ゼロ 圧縮空気加速 発電量増大  
垂直下方海水温度上昇ゼロ 人類削減 落差 大気圧同速度同容積仕事率先送り **船舶**  
海草 重力加速度加速 二酸化炭素排気ゼロ水  
工場電化全盛 **自動車 既存蒸気タービン発電**  
**電気駆動 既存世界 全面電化住宅** 全盛  
重力発電蓄電池駆動 既存火力原子力発電全廃 全廃地球 地球温暖化 人類

### Ta11.既存エンジンへの警鐘、タービン発電・重力発電

自動車 軽量物発電 運用  
安価電気駆動 軽量蒸気速度 横軸自動車 反転 永遠  
全動翼 太陽光加熱器熱製造 飛行機 液体酸素圧縮駆動  
**既存エンジン** 宇宙到達費用 **理論最良エンジン**  
船舶 電気+液体空気+過熱蒸気蒸気供給設備 大学  
**発電量** 既存 既存蒸気タービン発電  
容積圧縮仕事率 発電原価 後追いエンジン 発明阻止  
製造物全部 日帰り旅行 燃料費ゼロ空気圧縮  
**高校 静翼 燃費**

Ta10, Ta11は特定の出願人による重複した要約内容の特許から抽出された

※文字の大きさはトピックに対する関係の強さを表現している (上位5つの単語を赤色で表示している)

全特許データに対して各トピックの該当有無を計算し、各トピックが該当するデータのうち、“2014年以後”となる割合について、PLSAの結果とPCSAの結果を比較します

PLSAによるトピックとPCSAによるトピックのスコア(該当有無)を紐づけた特許データ

PLSAによる34個の全体トピック

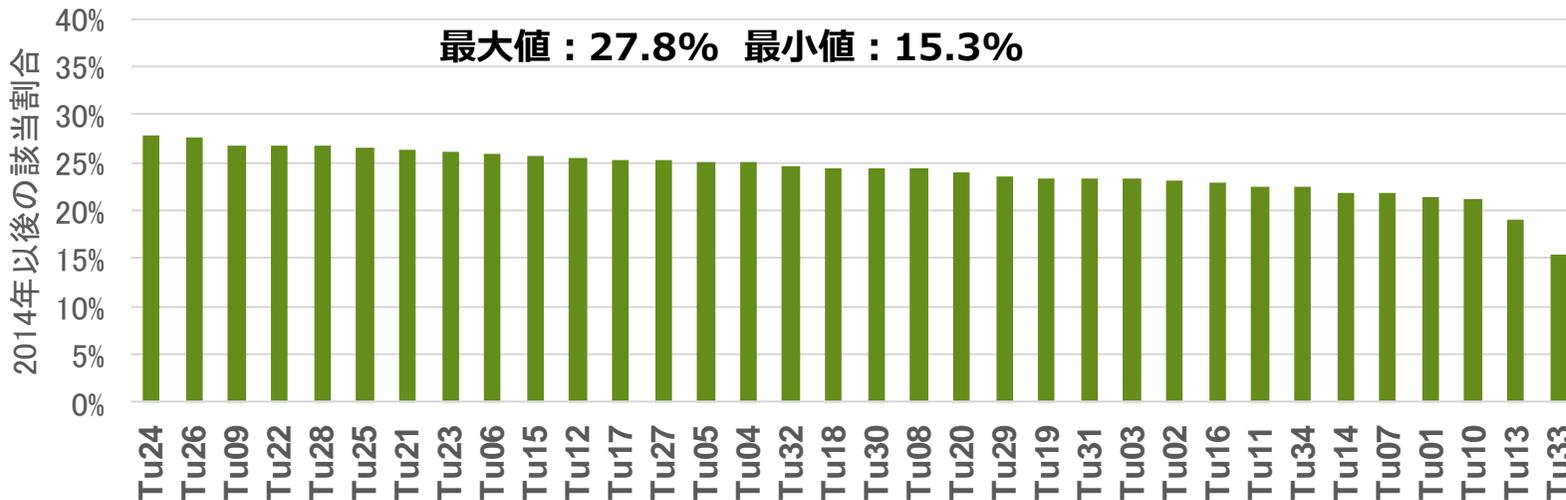
PCSAによる11個の特徴トピック

特許ID	出願番号	出願年	出願年グループ	トピック Tu01	トピック Tu02	...	トピック Tu34	トピック Ta01	トピック Ta02	...	トピック Ta11
1	特願2007-XXXX	2007	①2013年以前	1	1		0	0	1		1
2	特願2009-XXXX	2009	①2013年以前	0	1		1	0	0		1
3	特願2012-XXXX	2012	①2013年以前	0	1		1	0	1		0
4	特願2014-XXXX	2014	②2014年以後	1	0		0	1	0		0
...	...	...	...	...	...		...	...	...		...
26,419	特願2016-XXXX	2016	②2014年以後	1	0		1	1	0		0

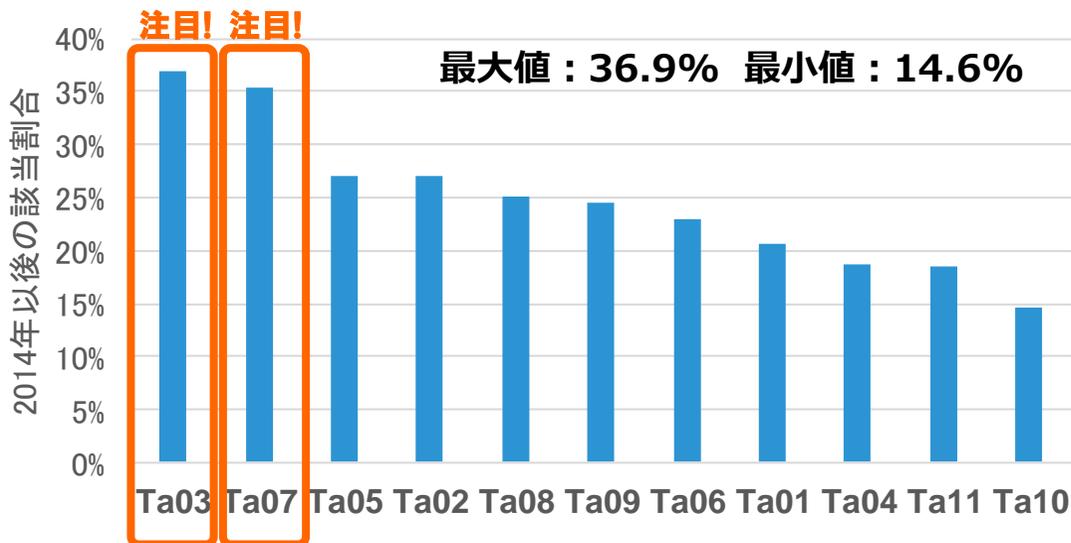
各トピックが該当するデータのうち、出願年が「②2014年以後」の割合を比較することで、2014年前後においてPCSAではどれくらい偏ったトピックを抽出できているか確認する

各トピックの2014年以後の割合は、PLSAではおおむね25%前後ですが、PCSAでは割合が高いものと低いものに偏っており、2014年前後に対して特徴的なトピックとなっています

PLSA  
による  
全体  
トピック



PCSA  
による  
特徴  
トピック



- それぞれ2014年以後の該当割合の高い順にトピックを並べている
- PLSAではおおむね25%前後のトピックが抽出されているが、PCSAのトピックでは割合が高いものから低いものまで抽出されており、最大値も高い
- (参考)元々の2014年以後データの割合は24.2%

2014年以後の割合が高いトピックとして、電力変換に関する制御技術(Ta03)、データに基づいた運転者のアシスト技術(Ta07)が近年上昇傾向にあることが分かります

## 2014年以後の割合がトップ2のトレンド



$$\text{※リフト値} = \frac{P(\text{出願年} | \text{トピック} Ty=1)}{P(\text{出願年})}$$

- リフト値は出願年とトピックの関係を示す指標
- トピック毎の各出願年の出願割合に対して、その出願年の出願割合で正規化した値であり、全体における各出願年の出願割合がそのトピックを条件にすることで何倍に変化するかを示す

## 該当特許の要約の例

### Ta03.交流・直流の変換

発明の名称	出願年
電気自動車	2016

#### 要約文 (抜粋)

車両の衝突時に平滑化コンデンサを放電する確実性を向上させる。衝突検知装置が、衝突を検知したときに第1信号と第1信号に続く第2信号を送信する。第1信号を受信したときに、インバータ制御回路がスイッチング素子駆動回路への電力の供給を停止する。第2信号を受信したときに、インバータ制御回路が平滑化コンデンサを放電する。

### Ta07.情報通信、検出判定システム

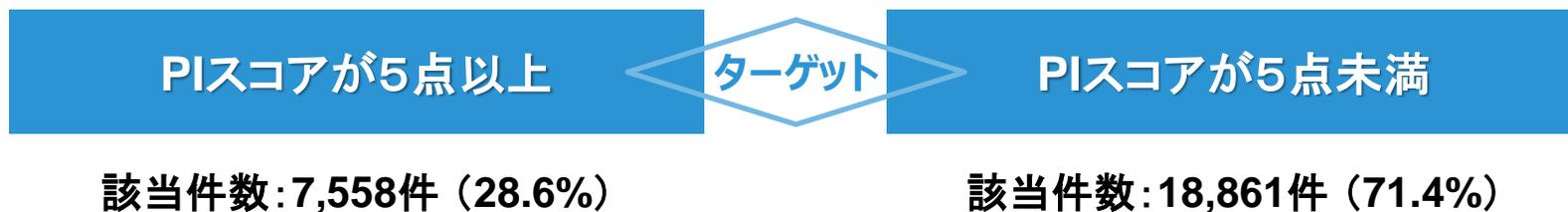
発明の名称	出願年
車両速度の制御方法	2016

#### 要約文 (抜粋)

経路及び交通状況に関する情報の応答として、自動車の速度を制御する方法に関するものである。本方法は、計画経路データ及び/または、繰返行程ロガーデータにより特定された想定経路に基づいて、最適な制動または加速点を決定し、最適な制動または加速点に基づいて、運転者に速度プロファイルを調整するためのサインを送る。

## 5. PCSAを適用した特許分析事例

### 5-4. PCSAによるPIスコアの特徴トピックの抽出



分析の  
ポイント

パテントスコアの高い有用な技術とは？

# PIスコアの特徴トピック10個の一覧

PIスコアで特徴を示すトピックは、エンジン駆動、電力変換、充電、二次電池の製造方法、冷却・加熱、検出判定、小型化・低コスト化など10個抽出されました

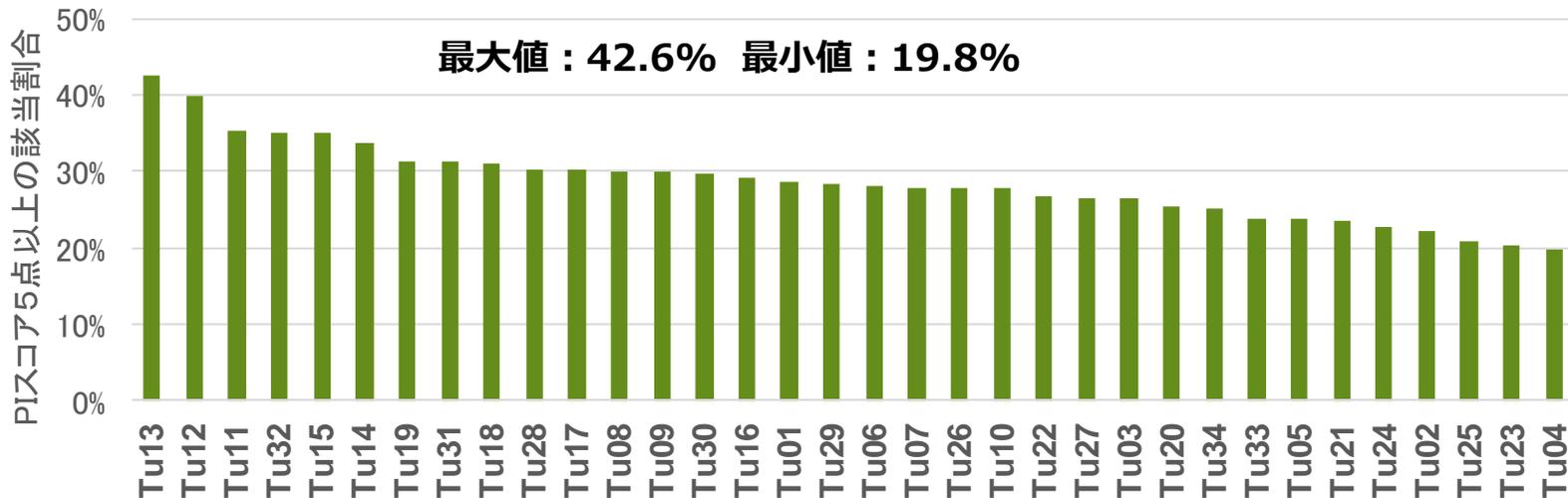


Tb09, Tb10は特定の出願人による重複した要約内容の特許から抽出された

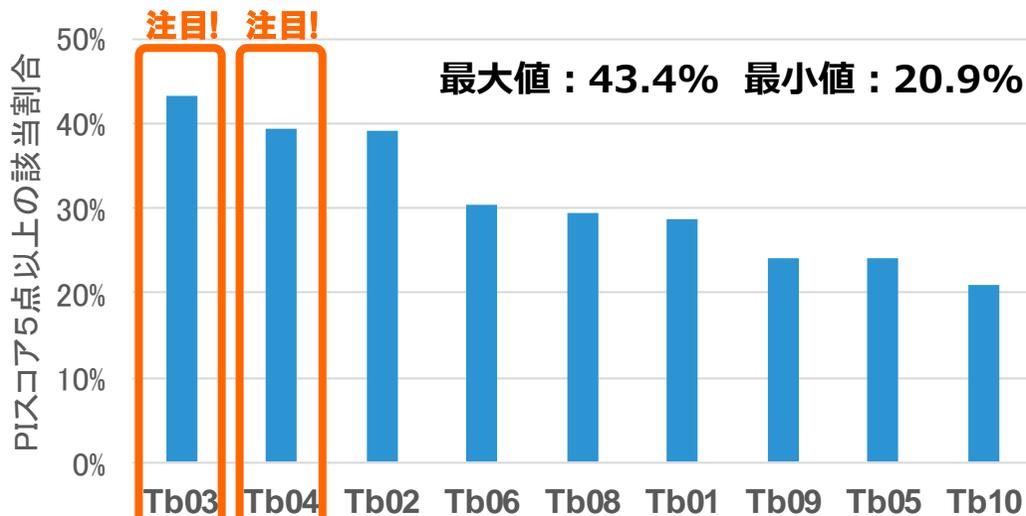
※文字の大きさはトピックに対する関係の強さを表現している（上位5つの単語を赤色で表示している）

各トピックでPIスコアが5点以上となる割合を比較すると、PLSAは平均的な割合のトピックが多く抽出されている一方で、PCSAはそうしたトピックは限定されています

PLSA  
による  
全体  
トピック



PCSA  
による  
特徴  
トピック

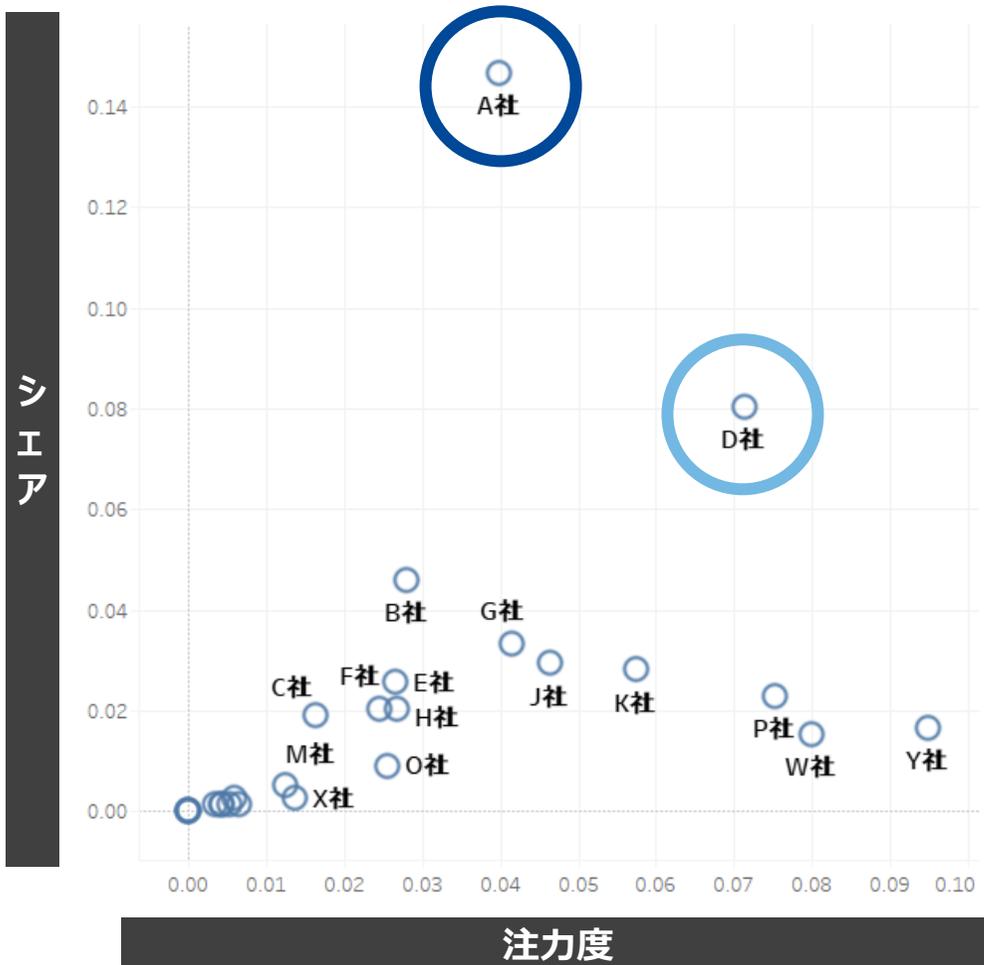


※Tb07は該当データがなかった

- それぞれPIスコア5点以上のデータの該当割合の高い順にトピックを並べている
- PLSAもPCSAも該当割合の最大値・最小値はほぼ変わらないが、PLSAでは平均的な割合のトピックが多く抽出され、PCSAではそれが限定的である
- (参考)元々のPIスコア5点以上データの割合は28.6%

電気自動車への充電に関するトピック(Tb03)はPIスコアが比較的高い傾向にありますが、メインプレーヤーは2社で、シェアがより高いA社と、注力度が高いD社となります

### 注力度とシェアを軸とした出願人のポジショニング



### 結果の解釈

- 最もシェアの高いのはA社で、2番手は少しギャップを開けてD社であるが、このトピックの領域ではこの2社がメインプレーヤーといえる
- D社はA社にシェアで劣るものの、注力度が高く、独自性の強い技術が開発されている可能性がある

### 注力度とシェア

#### ■ 注力度: $P(\text{トピック}T \mid \text{出願人}X)$

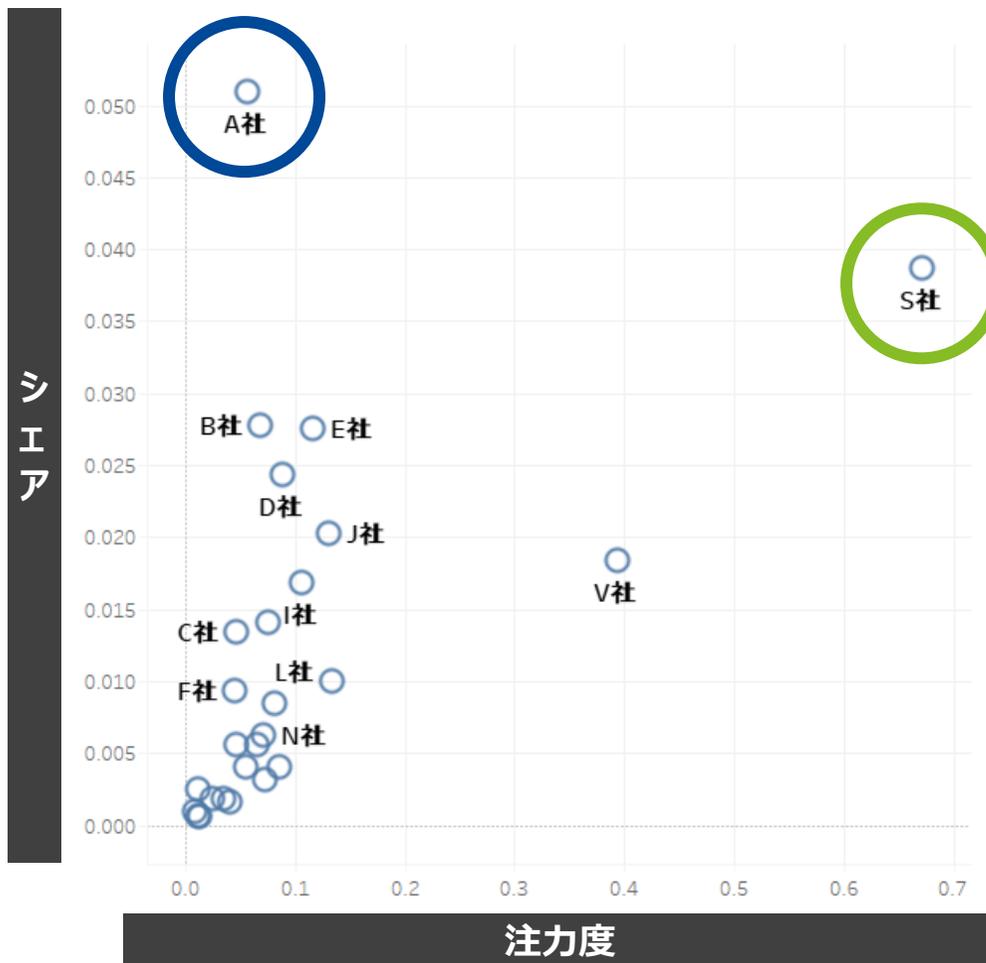
- 出願人Xの出願特許の中で、どれくらいの割合がそのトピックTに該当するものか、つまり出願人がどれくらいそのトピックに注力しているのかを示している

#### ■ シェア: $P(\text{出願人}X \mid \text{トピック}T)$

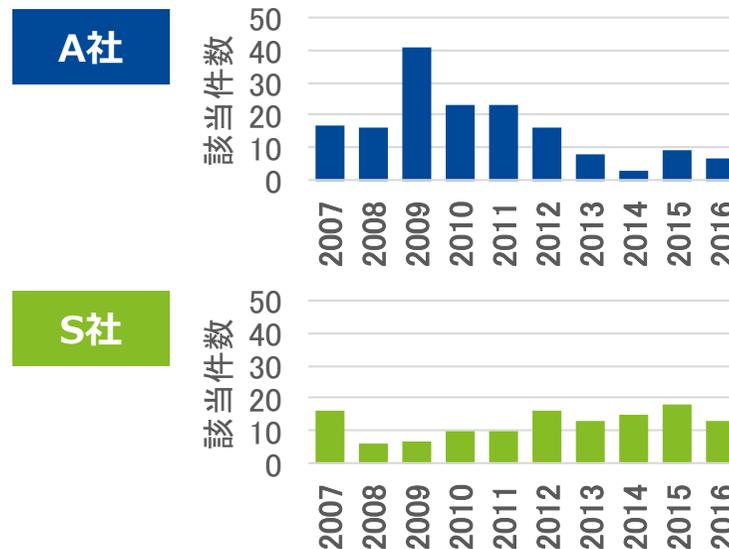
- トピックTが該当する特許の中で、どれくらいの割合がその出願人Xの出願によるものか、つまりトピックのなかでどれくらいその出願人が占めているのかを示している

二次電池の製造に関するトピック(Tb04)はPIスコアが高い傾向にありますが、メインプレーヤーはA社とS社の2社で、特に注力度が高いS社は出願数が徐々に増加しています

注力度とシェアを軸とした出願人のポジショニング



出願件数の推移

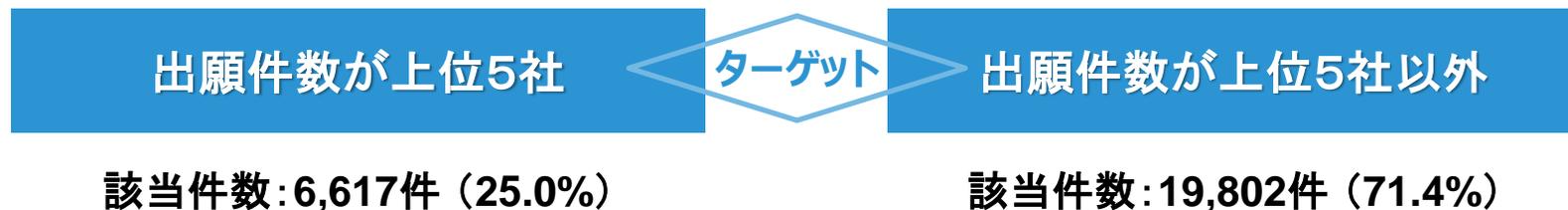


結果の解釈

- このトピックの領域ではA社とS社の2社がメインプレーヤーといえるが、特にS社は注力度が他社を大きく突き放して高く、独自性の強い技術が開発されている可能性がある
- 最近の出願件数では、A社は減少しているが、S社は徐々に増加しており、今後の出願動向には要注目である

## 5. PCSAを適用した特許分析事例

### 5-5. PCSAによる出願人の特徴トピックの抽出



分析の  
ポイント

出願件数の多い企業で牛耳られている技術とは？  
様々な出願人で棲み分けができていない技術とは？

※出願件数の多い上位5社とは、本資料ではA社、B社、C社、D社、E社に該当する

# 出願人の特徴トピック13個の一覧

出願人(上位5社か否か)で特徴を示すトピックは、エンジン駆動、モータ駆動、電力変換、充電、二次電池、温度制御、検出判定、操作性、小型化・低コスト化など13個抽出されました

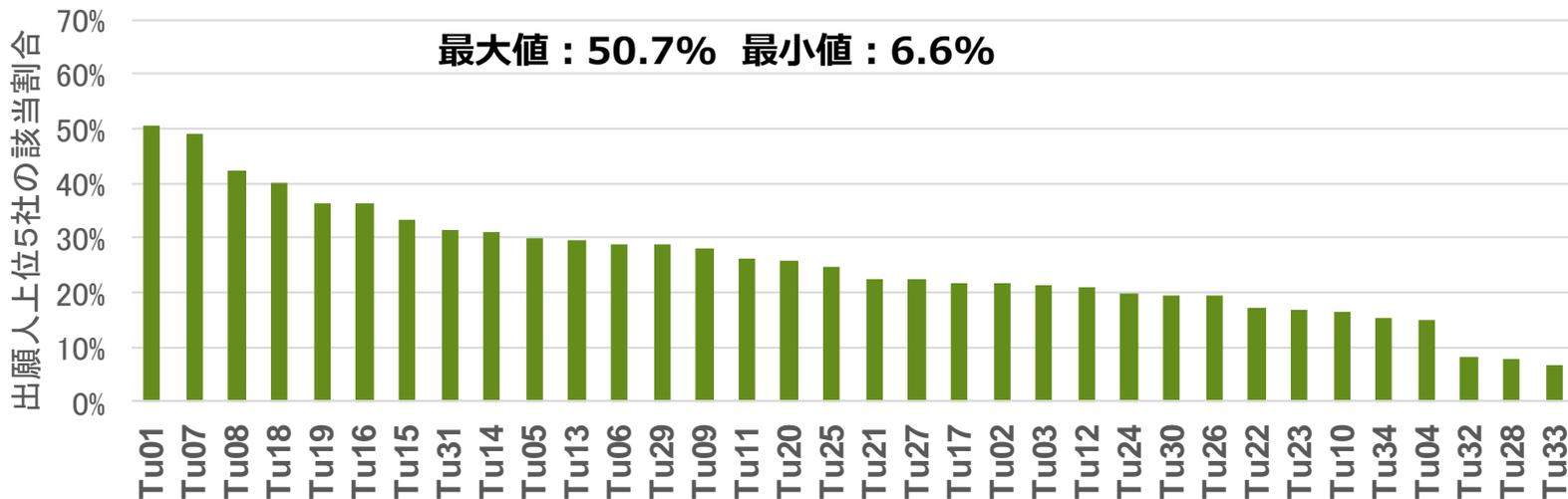
<h3>Tc01.エンジン駆動・動力伝達の制御</h3> <p>駆動輪動力 運転状態 開始 制御装置 電気式差動部 内燃機関停止 モータジェネレータ 燃費 モータハイブリッド車両 車両用駆動装置 車両用動力伝達装置 クラッチ エンジン 出力軸駆動 制御 トランスミッション バッテリ 駆動力変速部 構成 発電機 連結 走行</p>	<h3>Tc02.モータ駆動、ブレーキシステム</h3> <p>配置 移動 結合 ブレーキステータ 動作 駆動部 ロータ 生成方法 電気機械連結 制御 回転+できる システム位置 電気エネルギー 構成 モータ 回転車輪 自動車 回転軸 変換 駆動 支持 伝達 発電機 供給 作動</p>	<h3>Tc03.交流・直流の変換</h3> <p>遮断直流 コンバータオフ 放電 オン 検出 電気負荷 電力負荷 電圧 変換 制御 リール蓄電装置 インバータ バッテリ 制御装置 電気自動車 コンデンサ 供給 モータ 駆動 電力変換装置 電源装置 直流通力スイッチ 並列 電流 交流電力 作動</p>	<h3>Tc04.電気自動車への充電、給電装置</h3> <p>情報 コネクタ 充電ケーブル 蓄電池 制御 受電部 制御部 外部 走行 バッテリ 給電装置 制御装置 給電 充電+できる 充電 電源 電気自動車 蓄電装置 送信 供給 充電システム 構成 外部電源 電力 車両外部 充電スタンド 非接触放電 受電 電気機器</p>	<h3>Tc05.二次電池の構成・製造方法</h3> <p>強度 電極 製造方法 電極 単電池 表面 電池パック 配置 収容セパレータ 形成 リチウムイオン電池 正極 方法 負極 二次電池 積層 電池モジュール 構造 コネクタ 正極活物質 組電池 構成 バッテリ 負極活物質 電池特性 製造画 電解液 バッテリケース 集電体</p>
<h3>Tc06.空気や燃料電池などの温度制御</h3> <p>駆動制御 循環 冷却装置 バッテリ 排出 空気 水素 燃料内燃機関 発電機 エンジン 冷却 供給 冷却システム 燃料電池 熱 燃料電池システム 温度 排気ガス 作動 燃料電池車両 電力 変換 電気ヒータ 発電 加熱 モータ 車室内 冷媒</p>	<h3>Tc07.状態の検出と判定</h3> <p>送信 停止 情報 値方法 ECU バッテリープログラム 異常推定 受信 制御演算 判定 制御装置 検出 オン 取得 充電状態 閾値 比較 電圧 制御部 オフ 信号 記憶測定 温度 変化 有無</p>	<h3>Tc08.形成・配置</h3> <p>車体 方向 上方一対一体 連結位置 ケース一端 下方他端 收容 固定 コネクタ ハウジング 形成 配置 基板 構成 支持 接触 突出 対向 開口部 外部 端部 保持 端子 筐体 開口</p>	<h3>Tc09.スイッチの操作性</h3> <p>固定接点 導電ケース 検出 操作体 ストンプ 自動車 操作+できる 破損 車室内 電子機器 形成 スイッチ 下方 構成 スイッチ接点 装着 可動接点 入力装置 形成+できる 操作 安定 操作部 安価 接触</p>	<h3>Tc10.小型化・低コスト化・簡素化・信頼性向上</h3> <p>起因 実現 安定 製造コスト 確保 バッテリ コスト 安全 振動 低コスト 電気接続箱 検出+できる 構造 信頼性 電気自動車 効率 部品点数 確保+できる 影響 耐久性 劣化 小型化 技術 必要+ない 温度上昇 安価 損傷 モータ 構成 精度</p>
<h3>Tc11.既存エンジンへの警鐘・樹脂組成物の提供</h3> <p>電気特性 耐トランジエント性 組成物 電気部品 大学 製造 既存蒸気タービン発電 耐熱性 発電量 重量部後追いエンジン 発明阻止 高校 理論最良エンジン 自動車部品 ポリアリーレンスルフィド 既存エンジン 製造+できる 発電 含有 電気部品用途 全型 離型性 比重 大物質 地球最大 成形品 外観 機械的強度 エネルギー 成形品 溶融流動性 真空</p>	<h3>Tc12.重力発電の活用による地球温暖化防止</h3> <p>垂直下方 地球温暖化 全線地球温暖化 魚類 既存火力原子力発電全廃 圧縮空気加速 既存世界海水温度上昇ゼロ 発電量 燃料費ゼロ 重力加速度加速先送り 既存蒸気タービン発電 落差大気圧同速度同容積仕事率 安価 水 工場電化全盛 二酸化炭素排気ゼロ 人類絶滅 海草 発電量増大 船舶 全面電化 住宅全盛 電気駆動 重力発電蓄電池駆動 自動車 駆動タービン 重力発電運用</p>	<h3>Tc13.タービン発電の出力向上・燃費低減</h3> <p>蒸気速度 自動車 燃費 出力発電 横軸h歯車 既存蒸気タービン発電 反転 発電量 船舶 軽量物発電 空気圧縮 容積圧縮仕事率 宇宙到達費用マッハ狙い 液体酸素圧縮駆動 電気+液体空気+過熱蒸気 温熱供給設備D 容積 安価電気駆動 最大速度 海水 太陽光加熱器熱製造 運用 軽量蒸気速度 飛行機 製造物全部 改善 日帰り旅行発電原価 静翼 全動翼 既存 燃料費ゼロ 永遠</p>	<p>Tc11, Tc12, Tc13は特定の出願人による重複した要約内容の特許から抽出された</p>	

※文字の大きさはトピックに対する関係の強さを表現している (上位5つの単語を赤色で表示している)

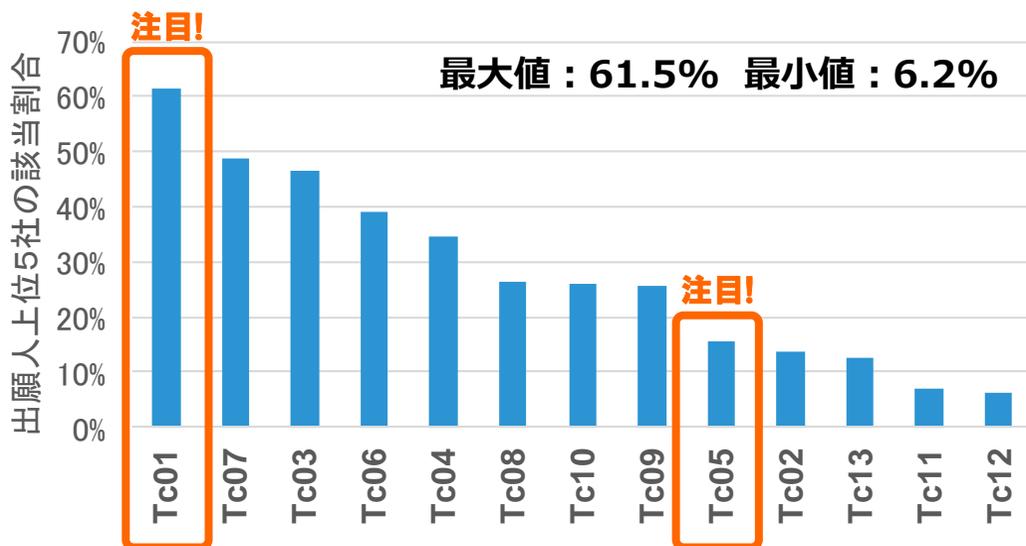
# PLSAとPCSAによるトピックの分布の比較

各トピックで出願件数が上位5社である割合を比較すると、どちらも高いものから低いものまで抽出されていますが、PCSAはその高低差が大きく、最大値も高い結果となっています

PLSA  
による  
全体  
トピック



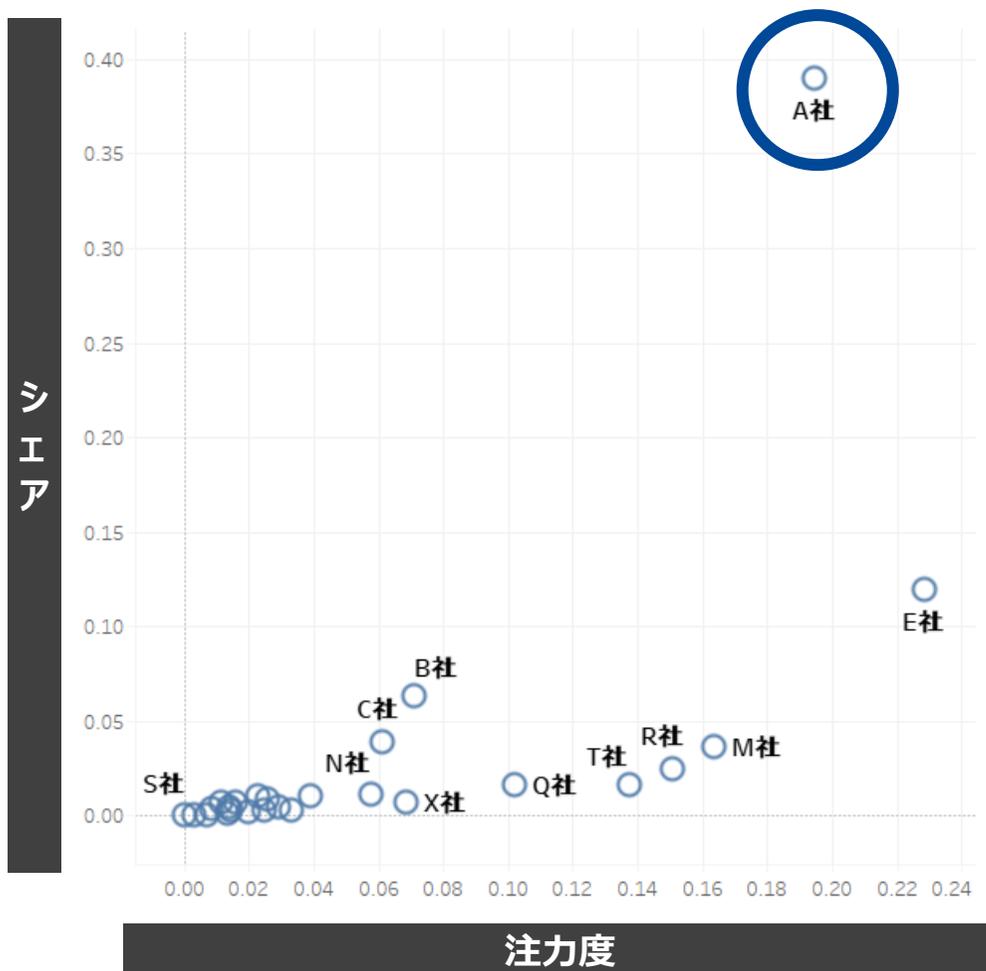
PCSA  
による  
特徴  
トピック



- それぞれ出願人上位5社のデータの該当割合の高い順にトピックを並べている
- PLSAもPCSAも割合の高いものから低いものまで抽出されているが、PCSAではその高低差が大きく、最大値も高く、そうしたトピックがPLSAよりも限定的に抽出されている
- (参考)元々の上位5社データの割合は25.0%

エンジン駆動や動力伝達の制御に関するトピック(Tc01)は上位5社の該当割合が高い傾向にありますが、実際にその5社のうちの1社がシェアを牛耳っており、一強状態にあります

### 注力度とシェアを軸とした出願人のポジショニング



### 結果の解釈

- A社が最もシェアが高く他社を大きく突き放しており、また注力度もとても高い水準にあり、一強状態にある
- E社は注力度では最も高く、シェアでは2番手だが、A社と大きくギャップを空けている

### 注力度とシェア(再掲)

#### ■ 注力度: $P(\text{トピック}T \mid \text{出願人}X)$

- 出願人Xの出願特許の中で、どれくらいの割合がそのトピックTに該当するものか、つまり出願人がどれくらいそのトピックに注力しているのかを示している

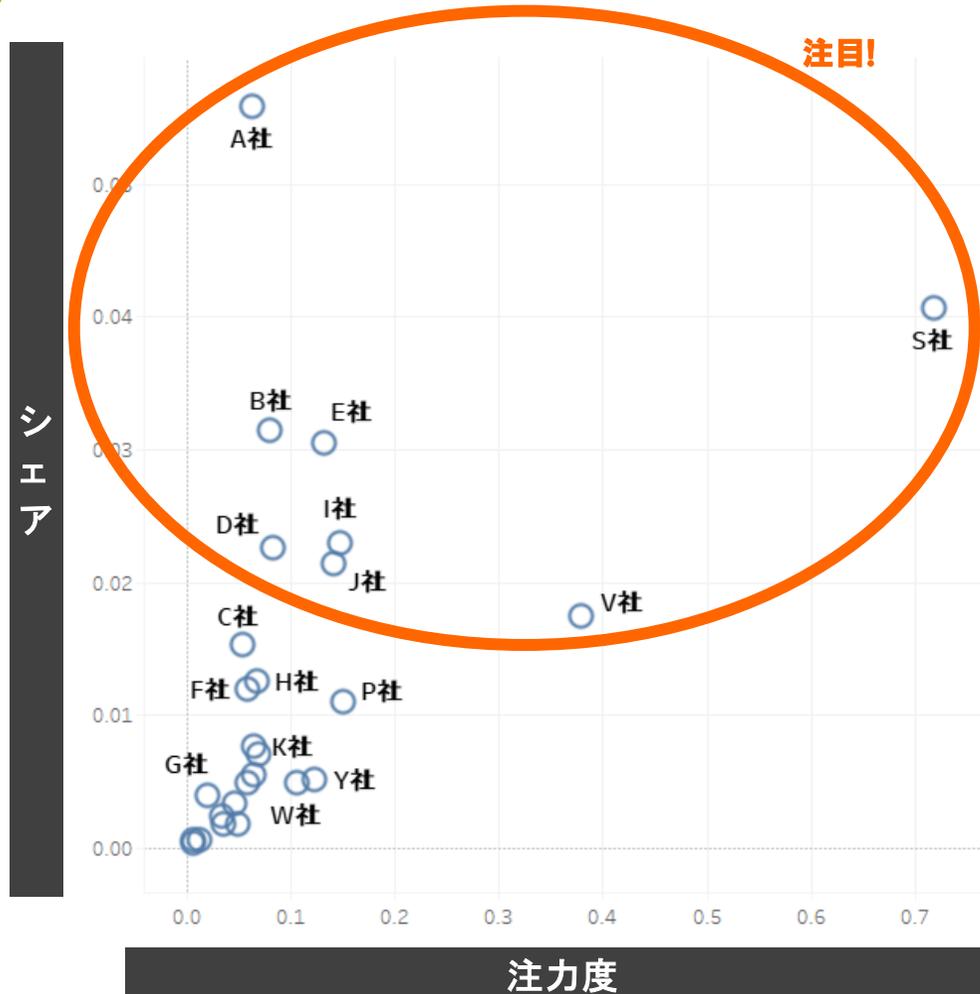
#### ■ シェア: $P(\text{出願人}X \mid \text{トピック}T)$

- トピックTが該当する特許の中で、どれくらいの割合がその出願人Xの出願によるものか、つまりトピックのなかでどれくらいその出願人が占めているのかを示している

二次電池に関するトピック(Tc05)は上位5社の該当割合が低い傾向にありますが、実際にプレーヤーが多く、5社以外の出願人も高シェア・高注力度の水準に位置しています

## 注力度とシェアを軸とした出願人のポジショニング

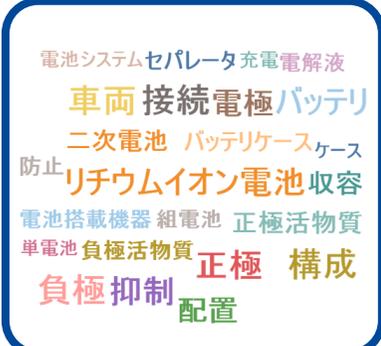
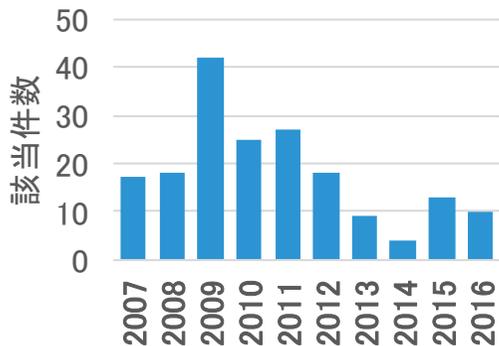
## 結果の解釈



- 最もシェアの高いのはA社で、次いでS社だが、他にも比較的シェアを保有している出願人が複数あり、プレーヤーが多い領域といえる
- 注力度ではS社が最も高く、他社を突き放しており、次いでV社の注力度も高い水準にある
- 巨大企業だけが占有するのではなく、各社が独自の技術を保有して棲み分けがされている可能性がある

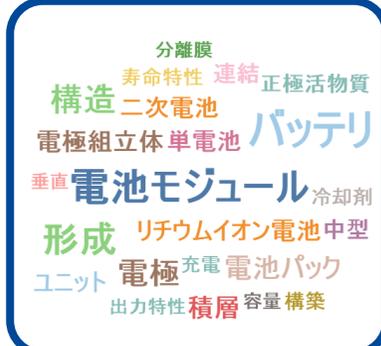
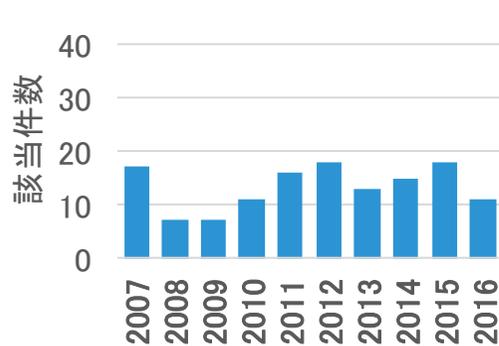
シェア1位のA社は2009年が出願ピークでリチウムイオン電池に関するものが多く、シェア2位注力度1位のS社は年10~20件の安定した出願で電池モジュールに関するものが多いです

A社(該当183件)の出願推移と要約キーワード



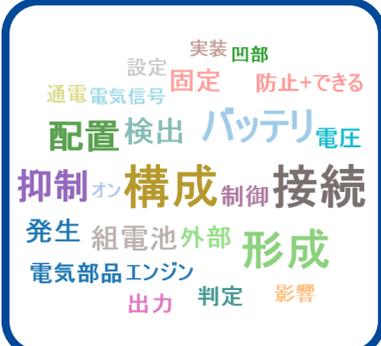
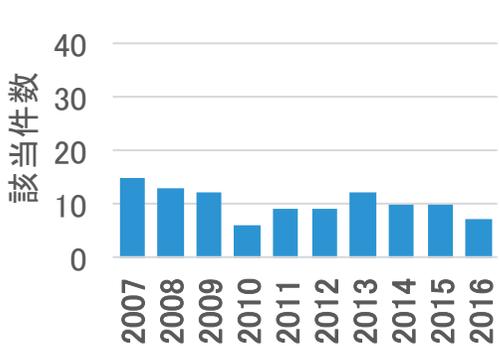
リチウムイオン電池や電極、接続・構成、收容ケースなど二次電池の全般的なキーワードが多い

S社(該当133件)の出願推移と要約キーワード



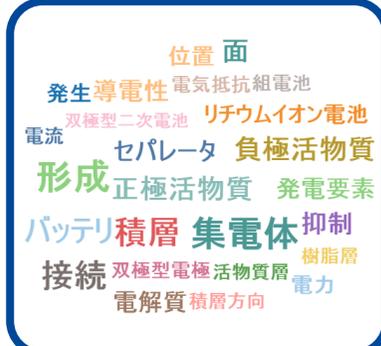
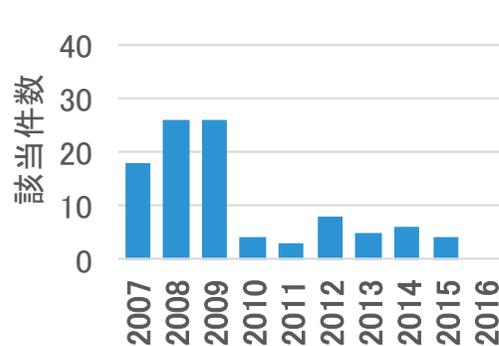
電池モジュールや電池パック、積層された電極組立体などに関するキーワードが多い

B社(該当103件)の出願推移と要約キーワード



構成・形成・接続や組電池などに関するキーワードが多い

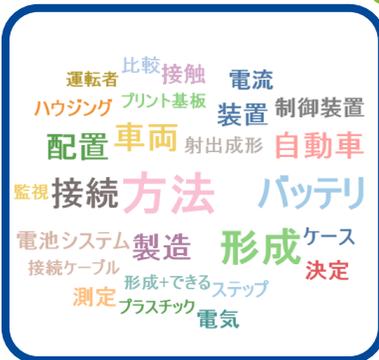
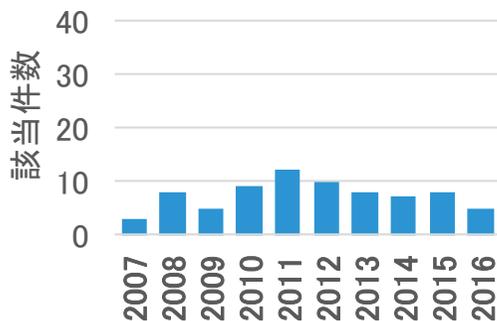
E社(該当100件)の出願推移と要約キーワード



集電体、電極活物質、積層、双極型電池などに関するキーワードが多い

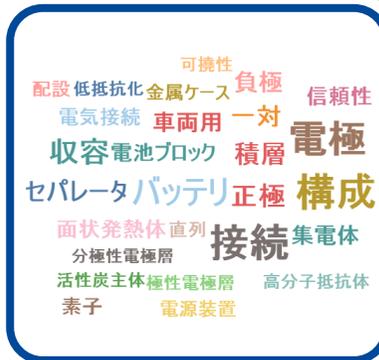
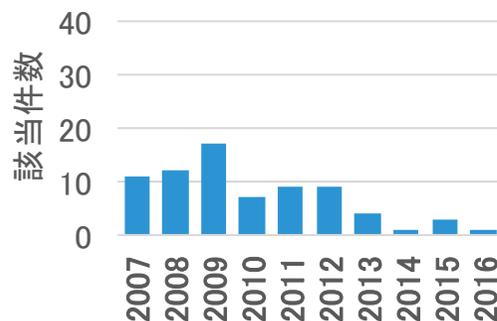
J社はシェア・注力度ともに高い水準ではないですが、直近で出願が急増しており、非水電解質や活物質に関するものが多く、注力度2位のV社は金属材料に関するものが多いです

### I社(該当75件)の出願推移と要約キーワード



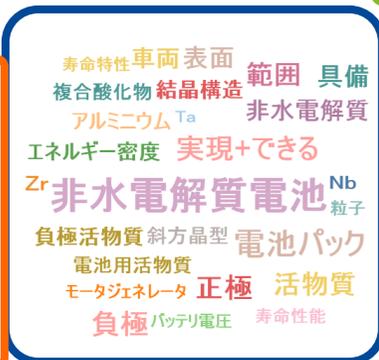
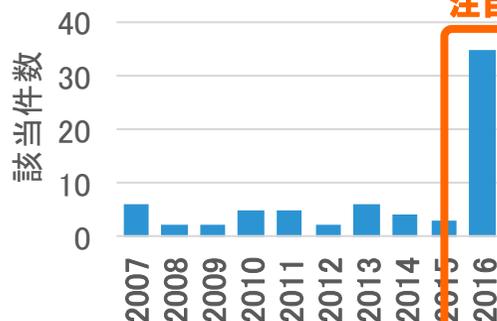
ケース、ハウジングなど筐体に関するキーワードが多い

### D社(該当74件)の出願推移と要約キーワード



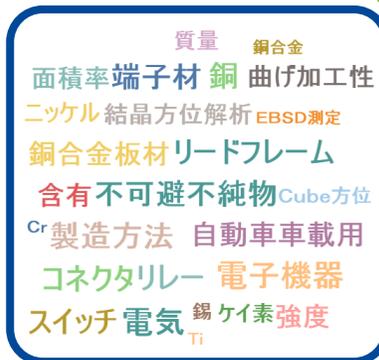
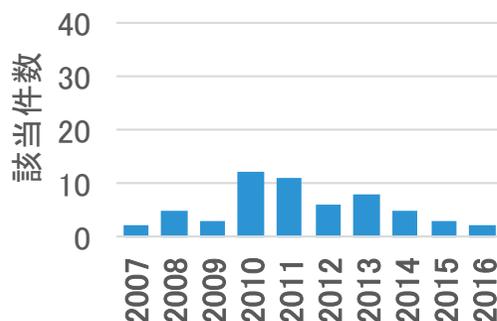
電極、セパレータ、構成などに関するキーワードが多い

### J社(該当70件)の出願推移と要約キーワード



非水電解質、活物質に関するキーワードが多い

### V社(該当57件)の出願推移と要約キーワード



電極の金属箔でも用いられる銅など金属材料に関するキーワードが多い

J社が2016年に出願した特許には、自動車への搭載を想定した寿命性能、放電性能、低温性能などに優れた非水電解質電池の技術があり、一部は登録済みとなっています

Tc05 × J社 × 2016年の該当特許35件から抜粋した6件の課題要約文

発明の名称	登録	PIスコア
電池パック及び自動車	済み	2
優れた急速充放電性能及び優れた繰り返し充放電特性を示すことができる活物質を含む電池パック及び自動車を提供する		

発明の名称	登録	PIスコア
組電池、電池パック、及び車両	済み	1
高いエネルギー密度及び高い電池電圧を示すことができ且つ寿命特性に優れた非水電解質電池を実現することができる電池用活物質を提供する		

発明の名称	登録	PIスコア
組電池、電池パック及び自動車	済み	1
高いエネルギー密度、高い電池電圧、及び優れた寿命特性を示すことができ、電圧管理を容易に行うことができる非水電解質電池を実現する		

発明の名称	登録	PIスコア
バイポーラ電池、電池パック及び車	済み	0
放電性能、低温性能、サイクル寿命性能及び高温貯蔵性能に優れた電気化学セルと、これを用いた電池パック及び車と、電気化学セルの製造方法とを提供する		

発明の名称	登録	PIスコア
固体電解質、リチウム電池、電池パック、及び車両	—	27
構成元素の化学的安定性を有しながら低コスト化が可能で、室温よりも低温でのリチウムイオン導電性が高い固体電解質、この固体電解質を含み、低温での放電レート性能が優れるリチウム電池、このリチウム電池を含む電池パック、及びこの電池パックが搭載された車両を提供する		

発明の名称	登録	PIスコア
電極、非水電解質電池、電池パック、及び車両	—	10
高温においても優れた寿命性能を示す非水電解質電池を実現できる電極、高温においても優れた寿命性能を示す非水電解質電池、この非水電解質電池を含む電池パック、及びこの非水電解質電池が搭載された車両を提供する		

※登録の確認は2018年9月7日時点

## 6. まとめ

## 6. まとめ

### 6-1. AIを応用した次世代テキスト分析技術 ～NomolyticsとPCSA～

AI技術を組み合わせて応用することで、膨大なテキストデータに潜む特徴や要因関係を可視化し、ビジネスアクションに有用なインサイトの獲得を期待できます

## Nomolytics®

テキストデータ全体を表すトピックを理解し、その平均的なトピックの特徴を様々な分析軸で探索できる

テキスト  
マイニング

PLSA

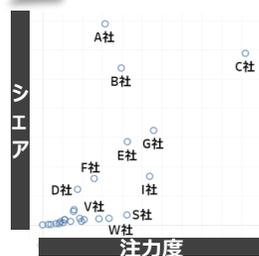
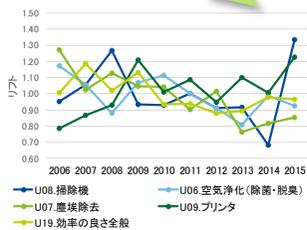
ベイジアン  
ネットワーク

単語抽出

トピック化

モデリング

全体を表現するトピックとは？



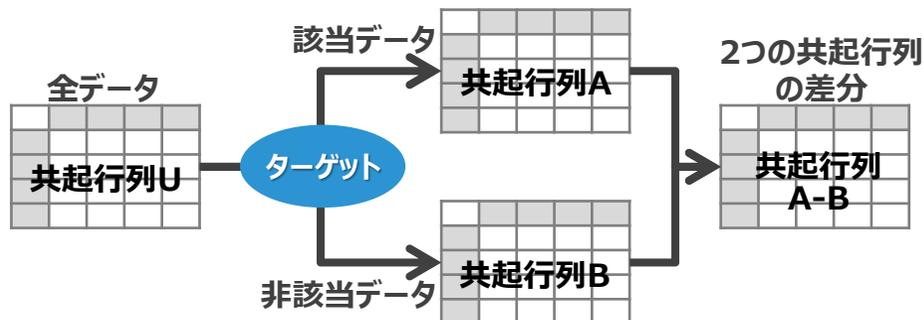
## PCSA®

探索したい特徴に特化したトピックを優先的に抽出し、より顕著な要因を深く分析してインサイトを得る

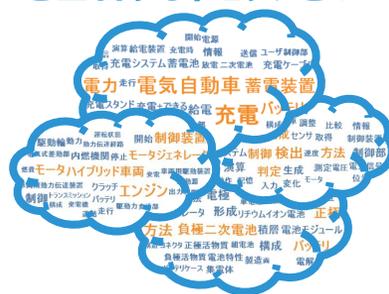
テキスト  
マイニング

差分の  
共起行列

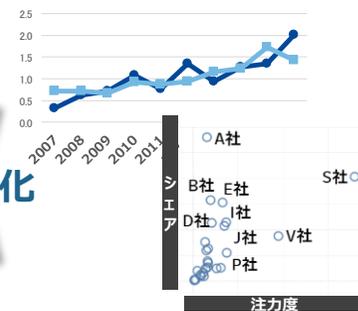
PLSA



ターゲットの該当有無を左右するトピックとは？



可視化



## NomolyticsもPCSAも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

- 製品やサービスのニーズをトピックで把握
- 顧客ターゲット別の関心トピックを把握
- 価値観に応じたマーケティングを検討



### アンケート

- 自由記述の内容をトピックで把握
- 他の設問と自由記述の関係を把握
- 顧客満足を生む施策を検討



### コールセンター履歴

- 問い合わせ内容をトピックで把握
- 各製品の問い合わせ特徴を把握
- 顧客離反を抑制する対策を検討



### 特許文書

- 特許文書の技術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
- 技術の差別化戦略や提携戦略を検討



### 営業日報

- 営業活動内容をトピックで把握
- 営業の成約確率を高める活動を把握
- 効果的な営業教育を検討



### 有価証券報告書

- 各企業の事業内容をトピックで把握
- 各種指標と事業トピックの関係を把握
- 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文の概要をトピックで把握
- 記述トピックに基づく学生の分類
- 学生の絞込みや面接の質問内容を検討



### 診療記録

- 診療記録、看護記録をトピックで把握
- 聞き取り内容と検査指標の関係を把握
- 定性情報も用いた診療支援を検討



### 問題発生レポート

- 不具合やヒヤリハットをトピックで把握
- 問題内容と作業環境の関係を把握
- 問題発生を抑制する環境改善を検討

## 6. まとめ

### 6-2. ビッグデータからインサイト獲得のためのススメ

# ビッグデータからインサイト獲得のためのダブルアプローチ

ビッグデータはそのままでは複雑で理解不能なので、まず理解できる形に抽象化して特徴を発見しますが、今度は抽象度が高く業務に活用できないので、その特徴の個別のデータに着目して再度具体化します

特徴を発見しやすくするために抽象化する (量的分析に軍配を上げ平均的な存在の間にある普遍的な特徴を得る)

## マクロ探索アプローチ

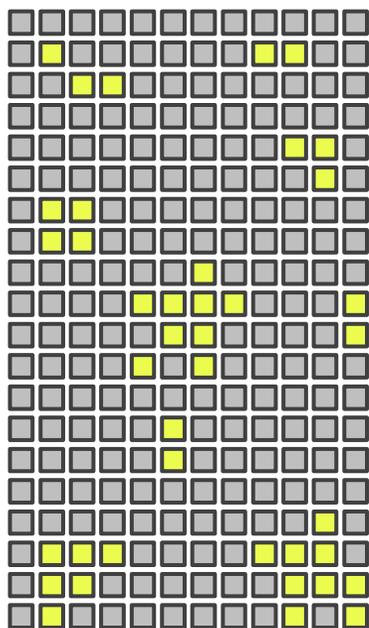
業務課題



確認したいデータの抽出

価値ある個別データの詳細調査と業務活用への昇華

膨大で複雑なビッグデータ

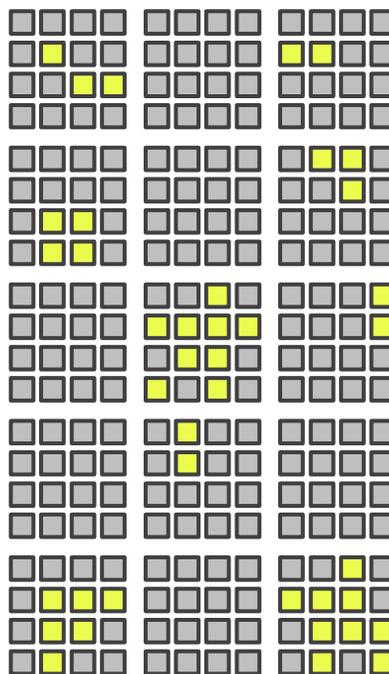


■ 価値あるデータ

特徴の似ているデータの分類

特徴トピックから個別データの絞り込み

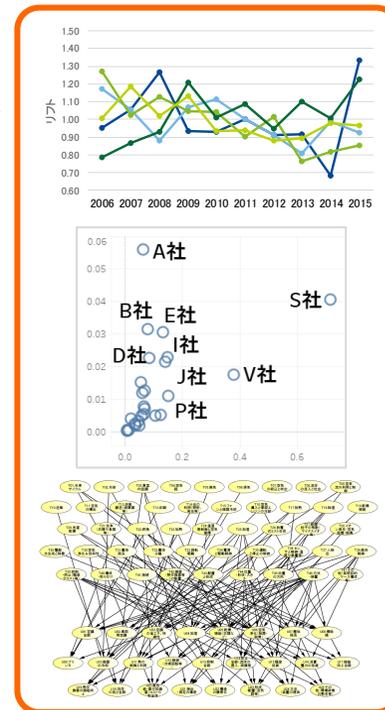
類型化



各トピックの特徴の可視化

注目すべきトピックや要因条件の発見

可視化・モデル化



## ミクロ探索アプローチ

価値ある個別のデータを発見するために具体化する (質的分析に軍配を上げ平均の中の個別の特性を確認する)

# ご清聴ありがとうございました

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](mailto: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。

<http://www.analyticsdlab.co.jp/>

株式会社アナリティクスデザインラボ

