



Analytics Design Lab

株式会社NTTデータ数理システム
ユーザーコンファレンス2018

電気自動車関連の特許文書データからAI技術で導くインサイト

確率的因果意味解析を用いた新たな分析軸の自動抽出とその利用

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

2018年11月22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月にデータ活用コンサルティングの新会社を設立しました

株式会社アナリティクスデザインラボ

企業におけるデータ活用を支援する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、というデータ活用するプロセスを企業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の課題解決を支援します。

| | |
|------|---|
| 設立 | 2017年6月1日 |
| 事業内容 | <ul style="list-style-type: none">● 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 |

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
- 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

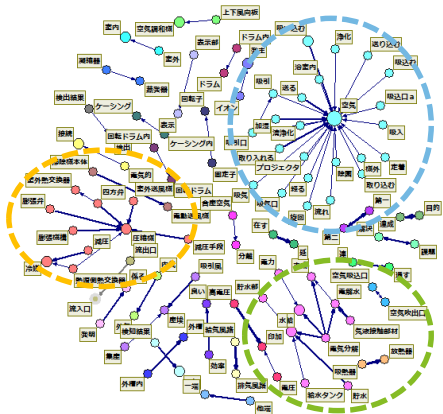
昨年の講演のおさらい

これまでの特許文書分析と課題

これまででは特許文書から抽出された単語やそれを手動でグルーピングしたカテゴリをベースに、全体の特徴を可視化していましたが、結果が複雑で解釈がしにくいものでした

これまでの特許文書分析のアウトプット例

単語共起ネットワークによる全体像把握

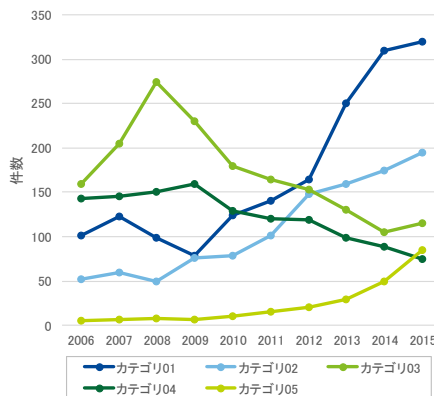


単語のグルーピング (カテゴリの作成)

例) 掃除機カテゴリ

| |
|----------|
| 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 |

カテゴリ別トレンド

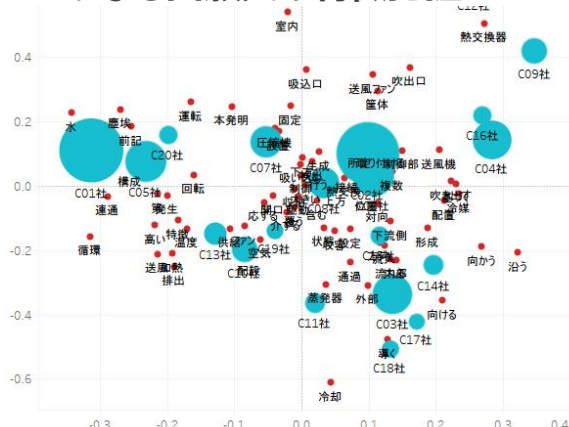


課題

単語ベースの分析では
複雑で考察しにくい

カテゴリの設定が主観的で
作業負荷も大きい

コレスポネンス分析による出願人の特徴把握



課題と解決手段のカテゴリのクロス集計による関係把握

| 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 | 10 | 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 |

課題と解決手段の統計的な
関係を分析していない

テキストマイニング技術に複数のAI技術を応用することで、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する技術を開発しました

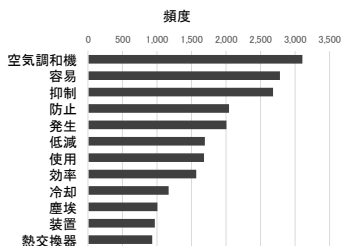
Nomolytics[®]: Narrative Orchestration Modeling Analytics

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

単語が出現する文脈を学習し、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

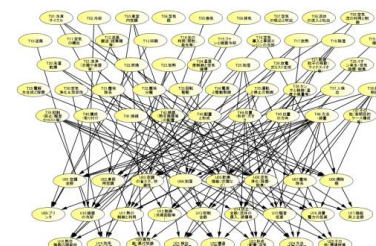


ベイジアンネットワーク

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

Nomolyticsの技術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

- 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ренд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

「風」「空気」に関する特許データ30,039件にNomolyticsを適用し、用途と技術のトピックの抽出、トピックのトレンドや出願人動向の可視化、用途と技術の関係分析を実施しました

トピックの抽出

テキストマイニングによる単語の抽出

| 単語 | 品詞 | 頻度 |
|-------|-----|-------|
| 空気調和機 | 名詞 | 3,106 |
| 空気 | 名詞 | 2,846 |
| 容易な | 形容詞 | 2,790 |
| 抑制する | 動詞 | 2,687 |
| 塵埃 | 名詞 | 1,687 |
| 分離する | 動詞 | 1,231 |
| ... | ... | ... |

PLSAの適用

【課題】の要約文から用途トピックを25個抽出

【解決手段】の要約文から技術トピックを47個抽出

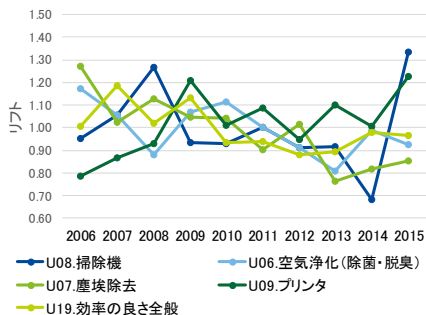
トピックのスコア計算

全特許データに対する各トピックの該当度を計算

| ID | 出願年 | 出願人 | 用途トピック1 | 用途トピック2 | 用途トピック* | 技術トピック1 | 技術トピック2 | 技術トピック* |
|----|------|-----|---------|---------|---------|---------|---------|---------|
| 1 | 2014 | A社 | 2.1 | 0.6 | ... | 1.5 | 5.0 | ... |
| 2 | 2013 | B社 | 0.3 | 3.4 | ... | 4.6 | 0.9 | ... |
| 3 | 2011 | C社 | 4.8 | 2.2 | ... | 2.7 | 1.1 | ... |
| n | ... | ... | ... | ... | ... | ... | ... | ... |

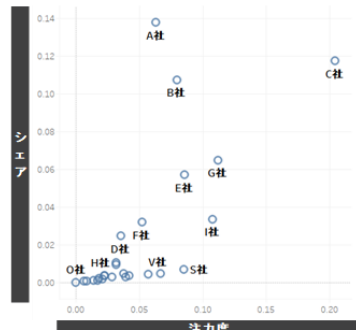
トレンドの可視化

各トピックのスコアを出願年で集計



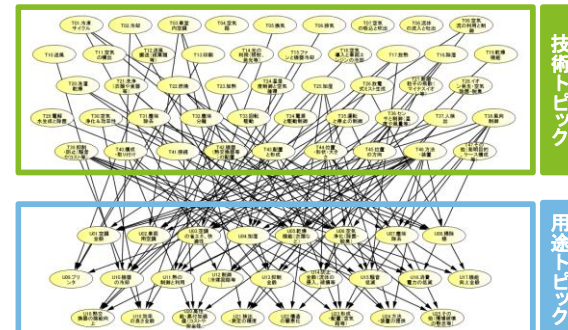
出願人動向の可視化

各トピックのスコアを出願人で集計して各社のポジションを可視化



用途と技術の関係分析

技術トピックに対する用途トピックの関係をページアンネットワークでモデル化し技術の新規用途を検討



昨年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、今年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て、より効果的な特徴の探索を行います

昨年は

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

適用
技術

Nomolytics

今年はい

属性の特徴を示す

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

適用
技術

PCSA(確率的因果意味解析)

PCSA: 確率的因果意味解析

PLSAは、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です

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することで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

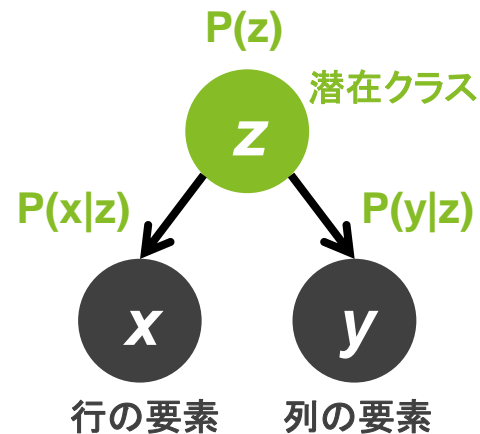
| 文書ID | 単語 1 | 単語 2 | 単語 3 | ... | 単語 5,014 | 単語 5,015 |
|------|------|------|------|-----|----------|----------|
| 1 | 0 | 0 | 1 | | 1 | 0 |
| 2 | 1 | 0 | 1 | | 0 | 1 |
| ... | | | | | | |



| 文書ID | トピック 1 | トピック 2 | ... | トピック 11 |
|------|--------|--------|-----|---------|
| 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

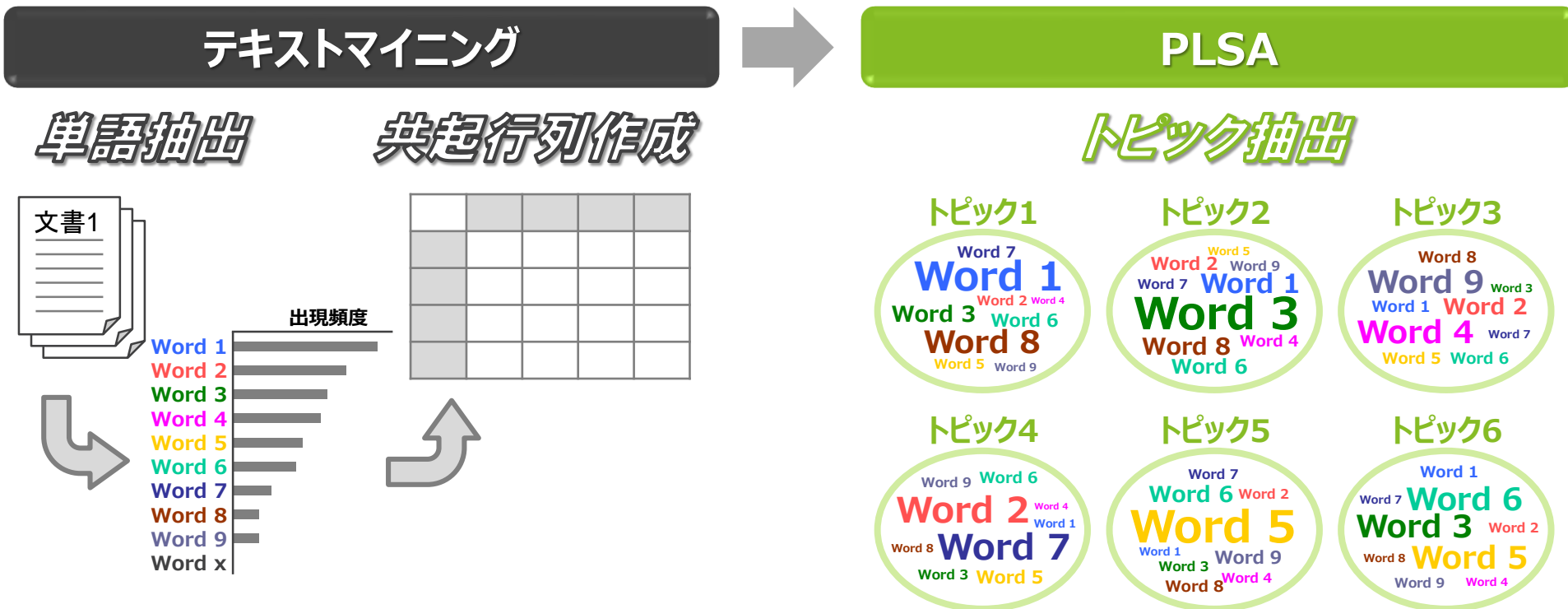
行の要素と列の要素を同時にクラスタリングできる

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

テキストマイニングで抽出された単語で構成された共起行列にPLSAを適用することで、単語をトピックに集約し、テキストデータの全体像をシンプルに把握します



テキストデータにテキストマイニングを実行して単語を抽出し、その単語の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、単語をトピックに集約する(使われ方の似ている単語をその重みと共にまとめる)

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

確率的因果意味解析 (PCSA: Probabilistic Causal Semantic Analysis)

テキストマイニング

PLSA

単語抽出



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

トピック抽出

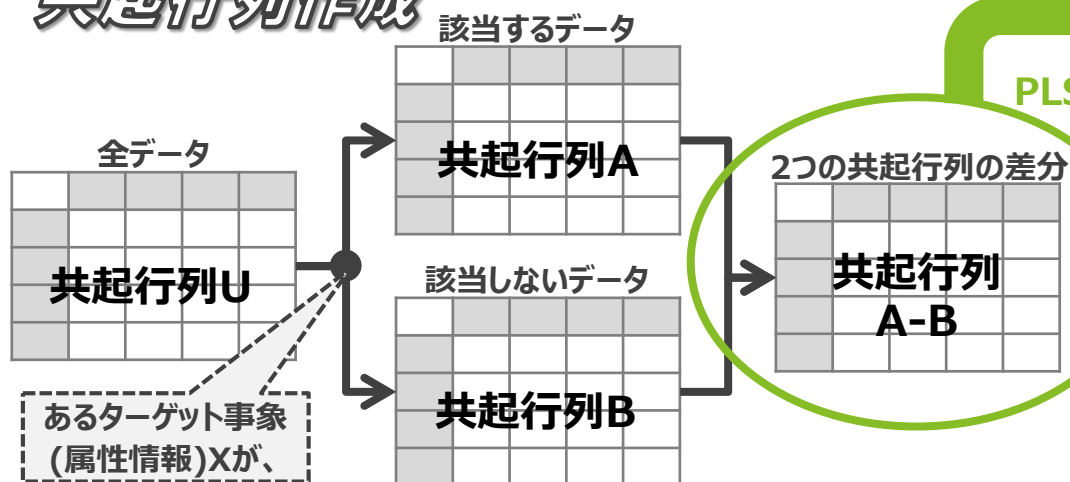


ターゲット事象Xの該当有無に影響を与える潜在トピックを優先的にテキスト情報から抽出できる



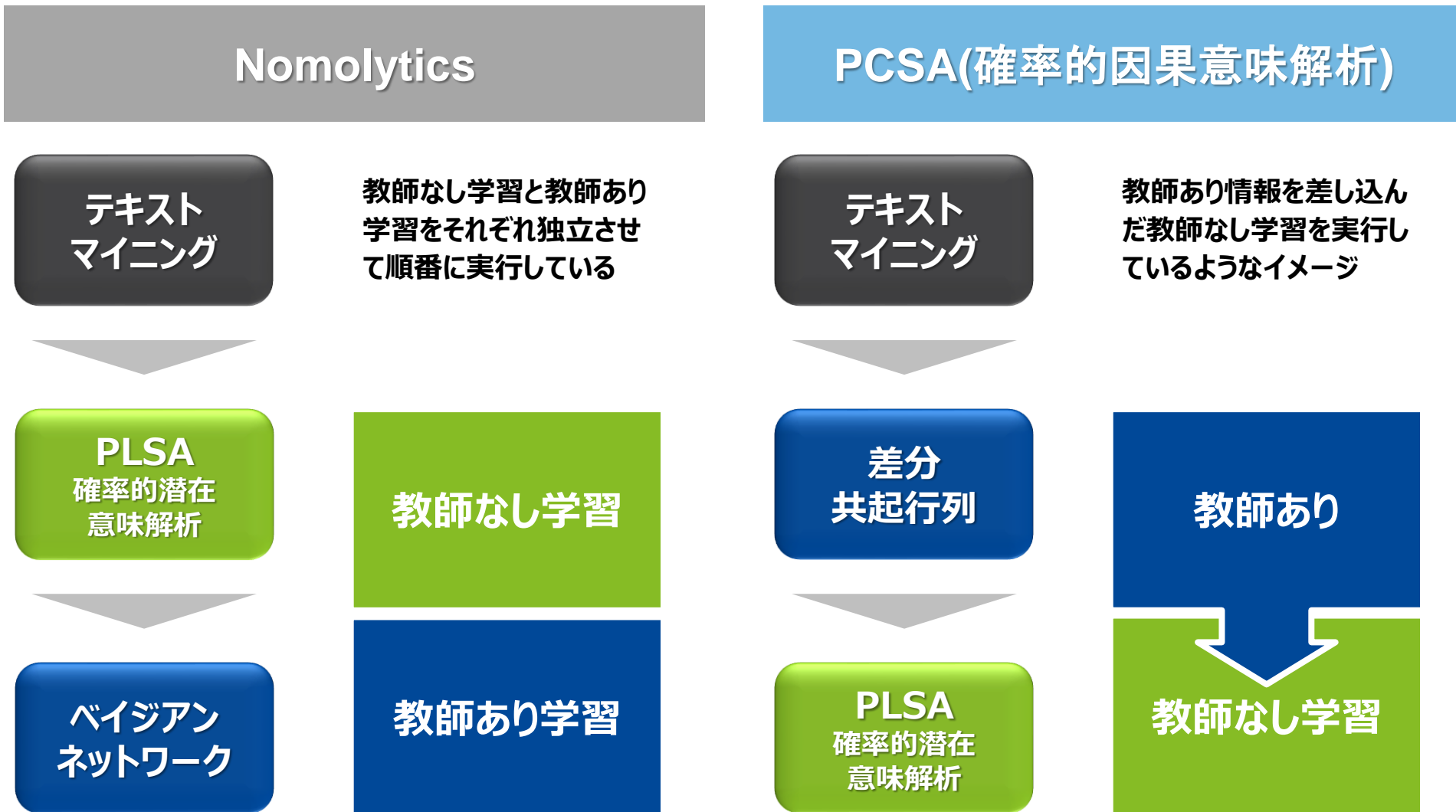
ターゲット事象 X

共起行列作成



【人工知能学会 2018年度全国大会優秀賞 受賞】

Nomolyticsでは教師なし学習(PLSA)を終了してから、その結果を使った教師あり学習をしています。PCSAでは教師の情報も考慮したような教師なし学習を実行しています。



PCSAの適用事例： 電気自動車関連の特許文書データの分析

※本事例は外部発表用に分析した事例であり、実際にサービスとして提供した分析の事例ではありません

「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、全体のトピックをPLSA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PCSAで抽出します

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
 - 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
 - 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)

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

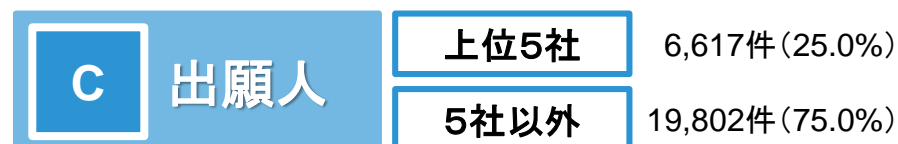
② PCSAで抽出する属性別の特徴トピック



▶ 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



▶ 出願量の多い企業とそうでない企業の棲み分けを把握する

全体のトピックの抽出

該当件数: 26,419件

トピック抽出手順①

テキストマイニングで要約文から名詞と係り受け表現を抽出し、名詞×係り受けで構成される共起行列を作成します

テキストマイニングの実行

Text Mining Studio

要約文に含まれる「名詞」と「係り受け表現」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 |
| バッテリー | 2,985 |
| 電力 | 2,828 |
| 駆動 | 2,496 |
| 方法 | 2,286 |
| 制御装置 | 2,092 |
| ... | ... |

名詞: 3,020語

| 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 |
| モータ-供給 | 270 |
| 並列-接続 | 267 |
| バッテリー-接続 | 252 |
| モータ-備える | 242 |
| 簡易-構成 | 228 |
| ... | ... |

係り受け: 2,128表現

- 係り受け表現は「名詞×動詞(サ変名詞含む)・形容詞・形容動詞」の係り受けペアを抽出
- 頻度20件以上を対象に抽出
- ノイズとなるような語はあらかじめ除外して抽出

共起行列の作成

名詞×係り受け表現の共起行列（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

| | 電力-供給 | 否-判定 | モータ-駆動 | バッテリー-充電 | 効率-良い | 供給-電力 | 電気自動車-提供 | 充電-行う | モータ-供給 | 並列-接続 |
|-------|-------|------|--------|----------|-------|-------|----------|-------|--------|-------|
| 構成 | 118 | 33 | 36 | 33 | 24 | 32 | 10 | 25 | 30 | 46 |
| 制御 | 268 | 73 | 108 | 85 | 12 | 115 | 2 | 41 | 74 | 40 |
| 配置 | 69 | 2 | 29 | 8 | 6 | 15 | 5 | 9 | 18 | 12 |
| モータ | 239 | 61 | 494 | 58 | 31 | 54 | 33 | 19 | 280 | 27 |
| 形成 | 31 | 4 | 20 | 8 | 1 | 7 | 0 | 5 | 4 | 12 |
| 供給 | 1,350 | 43 | 53 | 85 | 10 | 368 | 7 | 43 | 280 | 43 |
| 検出 | 134 | 99 | 56 | 36 | 4 | 44 | 8 | 19 | 33 | 27 |
| 電気自動車 | 193 | 73 | 79 | 129 | 59 | 53 | 289 | 114 | 54 | 13 |
| バッテリー | 337 | 87 | 60 | 529 | 44 | 81 | 20 | 72 | 106 | 70 |
| 電力 | 1,350 | 50 | 75 | 127 | 35 | 368 | 20 | 57 | 189 | 36 |

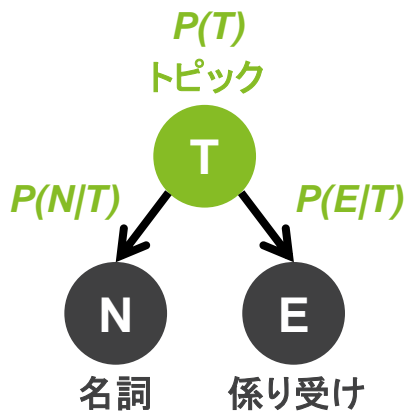
名詞: 3,020語 × 係り受け 2,128表現

作成した共起行列にPLSAを適用してトピックを抽出し、それぞれのトピックに所属する名詞と係り受けの所属確率からそのトピックの意味を解釈します

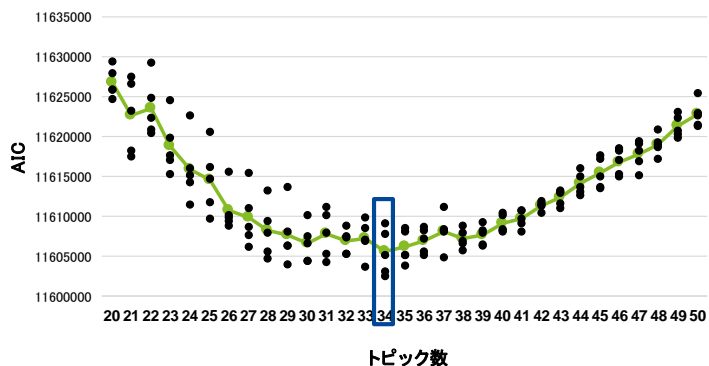
PLSAの実行



共起行列にPLSAを適用し、AICを評価基準とした最適なトピックを抽出する



- PLSAはVisual Mining Studioの二項ソフトクラスタリングを適用する
- PLSAはトピック数を指定する必要があり、初期値により解が異なる特性があるため、トピック数を20から50まで1刻みで変化させ、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



● それぞれの実行解のAIC
● 各トピック数における実行解の平均AIC

トピックの解釈

各トピックに所属する「名詞」と「係り受け」の所属確率からそのトピックの意味を解釈する

- PLSAでは以下の3つの確率がアウトプットとなる
 - $P(T)$ ・・・トピックの存在確率
 - $P(N|T)$ ・・・トピックにおける名詞の所属確率
 - $P(E|T)$ ・・・トピックにおける係り受けの所属確率
- $P(N|T)$ と $P(E|T)$ からトピックの意味を解釈する

(例) Tu13の解釈: 電気自動車の蓄電池充電

| $P(N U)$ | 名詞 | $P(E U)$ | 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 |
| 1.3% | 外部電源 | 1.1% | 充電-制御 |
| 1.2% | 充電+できる | 1.0% | 供給-電力 |
| 1.0% | 充電ケーブル | 0.9% | 蓄電装置-備える |
| 0.8% | 検出 | 0.9% | 電力-充電 |
| 0.7% | 情報 | 0.8% | 蓄電装置-提供 |
| ... | ... | ... | ... |

全体のトピック34個の一覧①

全体を集約するトピックでは、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、接続、小型化、安全性など34個抽出されました



全体のトピック34個の一覧②

全体を集約するトピックでは、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、接続、小型化、安全性など34個抽出されました

Tu21. 筐体

一方カバー 抽出 取容部 保持 モータ外部開口部
 筐体ケース内 形成 電子制御ユニット
 電気部品 開口 ハウジングケース 収容
 基板 制御回路配置 電気接続部 一体 挿入
 装着 コネクタ貫通孔 構成 対向 固定

Tu22. 表面の形成

一対外周面 被覆対向凹部 光 面
 周囲 外側一体 形状 先端発光素子
 導電性 突出 形成 接触 位置 表面
 構成 基板 貫通孔 本体 端部 部分
 配置 反対側 挿入電極 絶縁

Tu23. 位置とその移動

直交 許容形成 ロック 支持 係合 方向
 解除 保持 係 位置 回動 本体
 アウチュエータ ドア軸方向 回転 移動
 反対側 規制 固定 接触 移動+できる
 検出 電気信号

Tu24. 配置・位置・方向

供給 近傍領域 軸方向 モータ
 取容 ハウジング形成 配列 間隔 外側 離間位置
 対向隣接 配置 長手方向 方向
 流路 反対側 垂直 近接 平行
 構成 移動+できる 周囲端部 下方 一対

Tu25. 構成の方位

方向 車両前後 下面 配設
 電気自動車空間位置側面構成連結突出 配置
 車輪 車輪 後方 下方 開口 上方 一対 支持
 先端 形成 前方 収容 上面 バッテリー 電気掃除機
 固定 開口部ケース

Tu26. 構成

モータ制御+できる センサ 外部長さ
 自動車 生成 設置 構成 精度 確保 劣化 ハイブリッド車両 バッテリー
 電気自動車 充電 実用 制御装置 プロセッサ
 電力 依存 エネルギー貯蔵装置 接続 結合
 電気エネルギーシステム 電圧 供給 制御部 配置 コイル

Tu27. 接続

電源ケーブル 車体 固定 検出 保持
 位置 嵌合 電源プラグ 配線構成 基板 供給 接続部
 端子他端 コネクタ 一端 接続+できる
 配置 端部ケーブルバスバー 外部 形成 ワイヤハーネス 一対
 接地 電線 収容 回路

Tu28. 方法の提供

配置監視 段階 自動車 工程生成
 製造 エネルギー 電気機械調整 方法 バッテリー システム
 駆動 内燃機関 動作 存在 電気エネルギー 制御 モータ 分離 車両用 センサ

Tu29. 損傷や浸水など不具合の防止

耐久性 発明 衝撃 電気機器 不具合 静電気 構造 振動
 電気接続箱 製造コスト外部 確保+できる 影響
 損傷 侵入 電動パワーステアリング装置 未然 温度変化 信頼性 浸入
 外力 電気自動車 ノイズ 水 異音 安全 起因 衝突 破損 変動

Tu30. 小型化・簡素化・低コスト化など付加価値

リレー 大型化 自動車 リードフレーム 安価 小型 端子科 実現
 確保 低コスト 必要+ない コンパクト 製造コスト 車両用 灯具
 部品 点数 コスト 小型化 信頼性 構造
 簡素化 電気接続箱 耐久性 軽量化 製造方法
 コネクタ ワイヤハーネス 作業性 削減 強度

Tu31. 効率性・安全性の向上

燃費 温度上昇 電源システム 安定 実現
 構成 精度 確保 劣化 ハイブリッド車両 バッテリー
 安全 正確 電気自動車 モータ 効率
 制御手段 エネルギー効率 走行中 運転者 必要+ない 技術

Tu32. 既存エンジンへの警鐘・樹脂組成物の提供

含有成形品 耐熱性成分 組成物
 電気部品 自動車部品 既存蒸気タービン発電機 重量部
 発電量 理論最良エンジン 電気部品用途
 ポリアリレンスルフィド 耐衝撃性 後追いエンジン 発明阻止
 高校大学 機械的強度 既存エンジン 化合物 溶融流動性

Tu33. 重力発電の活用による地球温暖化防止

船舶 電気駆動 既存火力原子力発電全廃 圧縮空気加速 船舶
 落差燃料費ゼロ 垂直下方 全面電化住宅全盛
 工場電化全盛 駆動二酸化炭素排気ゼロ 全世界
 人類滅亡 大気圧同速度同容積仕事率 既存世界
 海水温度上昇ゼロ 先送り 既存蒸気タービン発電
 安価 重力加速度加速 重力発電運用用水 発電量増大
 タービン 重力発電蓄電池駆動 地球温暖化 自動車 発電量

Tu34. タービン発電の出力向上・燃費低減

反転 最大速度部水 発電原価 静翼 永遠運用改善
 横軸タービン 軽量蒸気速度 マッハ狙い 容積圧縮仕事率
 安価電気駆動 燃料費ゼロ 太陽光加熱器熱製造
 容積 電気+液体空気+過熱蒸気温熱供給設備
 宇宙到達費用 空気圧縮液体酸素圧縮駆動 発電量
 軽量物発電日帰り旅行 飛行機 製造物全部
 燃費 既存蒸気タービン発電 自動車 既存 全動翼 船舶
 蒸気速度 出力発電

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

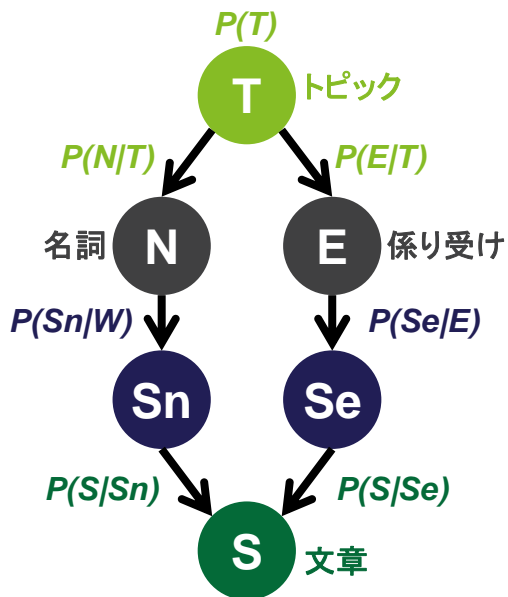
※文字の大きさはトピックに対する関係の強さを表現している (上位5つの単語を赤色で表示している)

各データに対するトピックの該当有無の計算

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した

| | |
|--------------|-----------------------|
| 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



文章を名詞で定義される文章Sn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

| |
|--|
| 名詞 N_i で定義される文章 Sn_h |
| $Sn_h = \{N_1, N_2, \dots, N_i\}$ |
| トピック T_k を条件とした文章 Sn_h の出現確率 |
| $P(Sn_h T_k) = \sum_i P(Sn_h N_i)P(N_i T_k)$ |
| 名詞 N_i が出現する中で文章 Sn_h が出現する確率(N_i の出現文章数の逆数) |
| $P(Sn_h N_i) = 1/n(N_i)$ |
| 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 |
| ※ $P(S_h Sn_h)$ と $P(S_h Se_h)$ はともに1/2とする |
| $P(S_h T_k) = P(S_h Sn_h)P(Sn_h T_k) + P(S_h Se_h)P(Se_h T_k)$ |
| 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 | ... | T34 |
|------|------|-----|-----|-----|-----|-----|
| 1 | 1 | 3.1 | 0.9 | 2.0 | | 3.5 |
| 1 | 2 | 1.4 | 0.2 | 5.5 | | 8.4 |
| 2 | 1 | 0.8 | 5.8 | 1.3 | | 2.9 |
| 2 | 2 | 1.2 | 3.2 | 1.7 | | 4.0 |
| 2 | 3 | 0.6 | 1.8 | 2.6 | | 9.6 |
| ... | | | | | | |

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 | ... | T34 |
|------|-----|-----|-----|-----|-----|
| 1 | 3.1 | 0.9 | 5.5 | | 8.4 |
| 2 | 1.2 | 5.8 | 2.6 | | 9.6 |
| ... | | | | | |

③閾値を設定してフラグに変換する

※閾値は5に設定した

| 特許ID | T01 | T02 | T03 | ... | T34 |
|------|-----|-----|-----|-----|-----|
| 1 | 0 | 0 | 1 | | 1 |
| 2 | 0 | 1 | 0 | | 1 |
| ... | | | | | |

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集計・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 | 要約 | 出願年 | 出願人 | PIスコア | トピック Tu01 | トピック Tu02 | ... | トピック Tu34 |
|--------|-------------|------------------|------|-----|-------|--------------|--------------|-----|--------------|
| 1 | 特願2007-XXXX | 【課題】電気式変速操作装... | 2007 | A社 | 0 | 1 | 1 | | 0 |
| 2 | 特願2009-XXXX | 【課題】従来の電気自動車... | 2009 | B社 | 1 | 0 | 1 | | 1 |
| 3 | 特願2012-XXXX | エンジンのための方法及び... | 2012 | C社 | 13 | 0 | 1 | | 1 |
| 4 | 特願2013-XXXX | 【課題】駐車場に設置された... | 2013 | D社 | 7 | 1 | 0 | | 0 |
| ... | ... | | ... | ... | | ... | ... | | ... |
| 26,419 | 特願2016-XXXX | 充電ステーションが電気エネ... | 2016 | X社 | 0 | 1 | 0 | | 1 |



トピック×属性の様々な集計・分析が可能に

PCSAの適用: 出願年の特徴トピックの抽出

2014年以後

該当件数: 6,405件 (24.2%)

2013年以前

該当件数: 20,014件 (75.8%)

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?

特徴トピックの抽出手順

①2014年以後データ、②2013年以前データで作成した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のあるトピックを優先的に抽出します

2つの共起行列の作成

差分の共起行列の作成

PLSAの適用

①出願年が2014年以後のデータ

6,405件(文章数:10,723件)

| | 電力-供給 | 否-判定 | モーター-駆動 | バッテリー-充電 | 効率-良い |
|------|-------|------|---------|----------|-------|
| 構成 | 21 | 3 | 9 | 8 | 6 |
| 制御 | 60 | 24 | 28 | 19 | 3 |
| 配置 | 13 | 1 | 6 | 0 | 2 |
| モーター | 47 | 19 | 138 | 12 | 5 |
| 形成 | 7 | 2 | 7 | 0 | 0 |

②出願年が2013年以前のデータ

20,014件(文章数:33,113件)

| | 電力-供給 | 否-判定 | モーター-駆動 | バッテリー-充電 | 効率-良い |
|------|-------|------|---------|----------|-------|
| 構成 | 97 | 30 | 27 | 25 | 18 |
| 制御 | 208 | 49 | 80 | 66 | 9 |
| 配置 | 56 | 1 | 23 | 8 | 4 |
| モーター | 192 | 42 | 356 | 46 | 26 |
| 形成 | 24 | 2 | 13 | 8 | 1 |

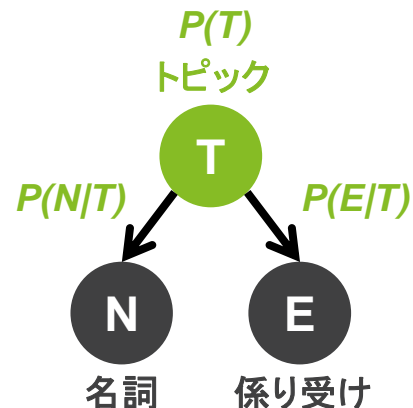
①2014年以後の共起行列と
②2013年以前の共起行列の
差の絶対値を計算した共起行列を作成する

| ①一重み調整済みの② |

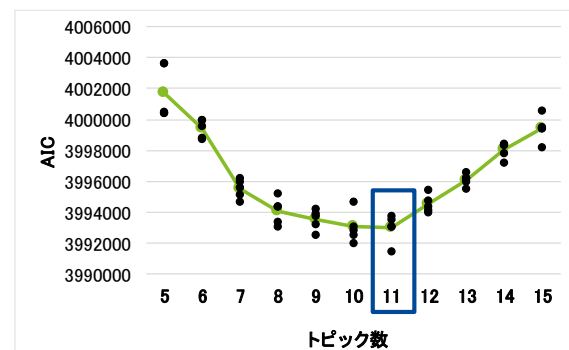
| | 電力-供給 | 否-判定 | モーター-駆動 | バッテリー-充電 | 効率-良い |
|------|-------|------|---------|----------|-------|
| 構成 | 10.4 | 6.7 | 0.3 | 0.1 | 0.2 |
| 制御 | 7.4 | 8.1 | 2.1 | 2.4 | 0.1 |
| 配置 | 5.1 | 0.7 | 1.4 | 2.6 | 0.7 |
| モーター | 15.2 | 5.4 | 22.7 | 2.9 | 3.4 |
| 形成 | 0.8 | 1.4 | 2.8 | 2.6 | 0.3 |

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、②2013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、①2014年以後の共起行列との差の絶対値を計算する

差分の共起行列にPLSAを適用する



AICを評価基準とした最適なトピックを抽出する



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
トランスミッション制御手段駆動運転者 **モータ**
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
収容移動+できる 電気機械 駆動部 **回転**
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 **電力** 給電装置 コネクタ 充電+できる
充電システム **電気自動車** 蓄電装置
充電スタンド外部電源 **充電** **バッテリー**
給電開始供給情報 **充電** 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 **駆動** 加熱 構成 冷却水 **発電**
エンジン 電気エネルギー **供給** **発電機**
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 人類滅亡 発電量増大
海草 重力加速度加速 二酸化炭素排気ゼロ 水
落差 大気圧同速度同容積仕事率先送り **船舶**
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 全盛
Top: 1 全盛
0.039

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

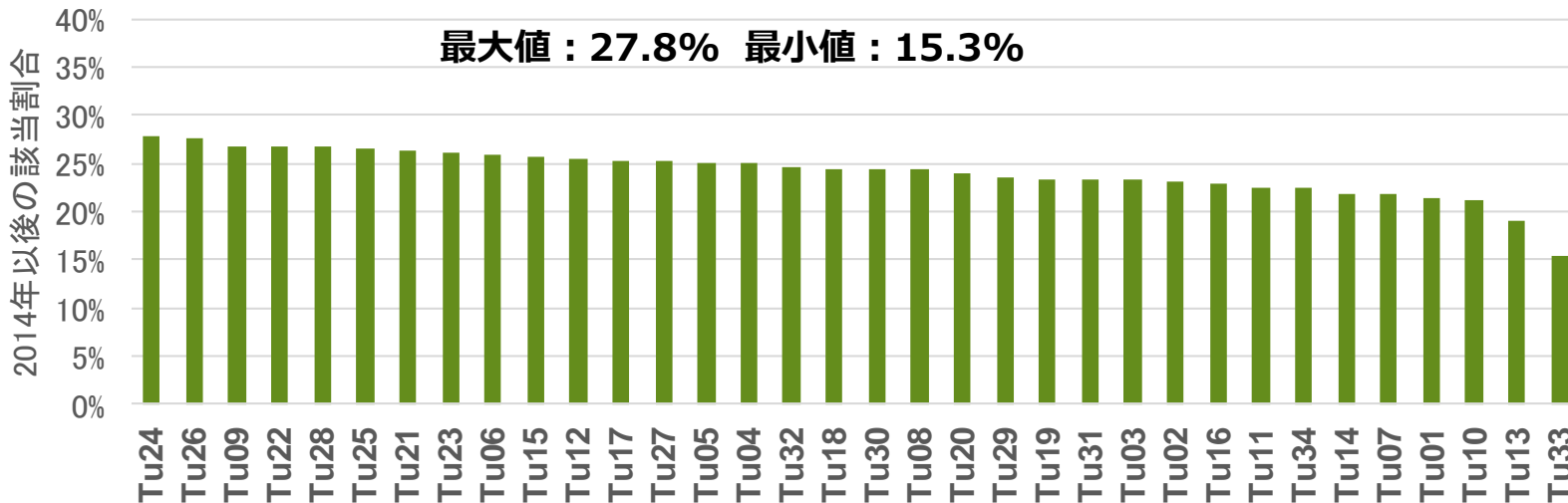
Ta10, Ta11は特定の出願人による重複した要約内容の特許から抽出された

※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)

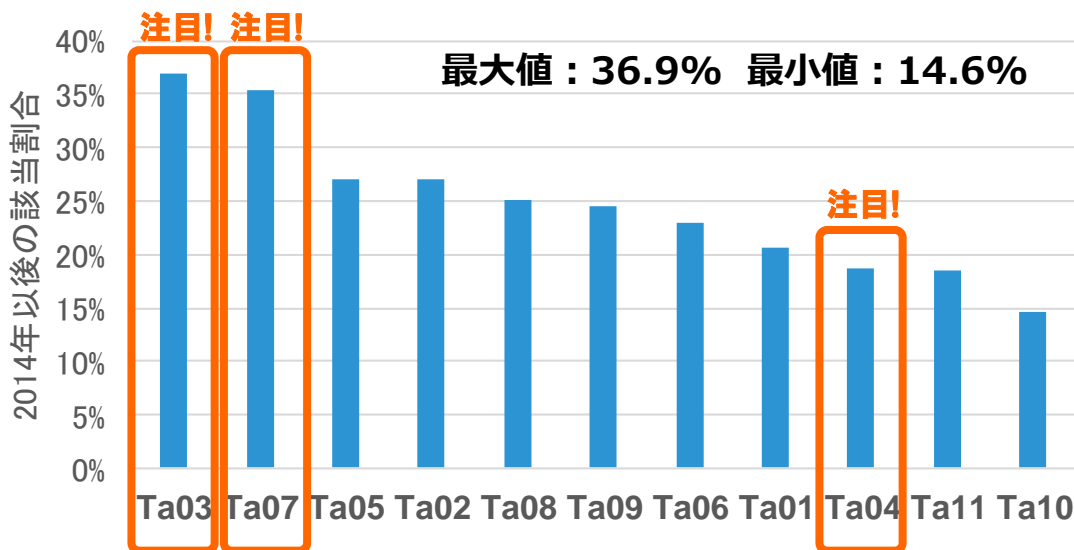
全体のトピックとの属性分布の比較

トピックごとに2014年以後データの該当割合を比較すると、全体ではおおむね25%前後ですが、PCSAのトピックでは高いものから低いものまで(特に高いものが)抽出されています

全体のトピック



PCSA適用の特徴トピック

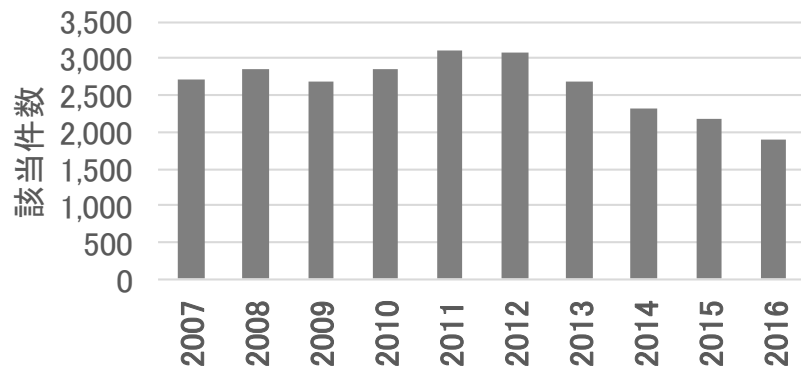


- それぞれ2014年以後のデータの該当割合の高い順にトピックを並べている
- 全体のトピックではおおむね25%前後のトピックが抽出されているが、PCSAのトピックでは割合が高いものから低いものまで抽出されており、最大値も高い
- 元々の2014年以後データの割合は24.2%

2014年以後の該当割合が低いTa04は2011年が出願のピークとなっており、2014年以後の該当割合が高いTa03、Ta07は近年出願が多くなっています

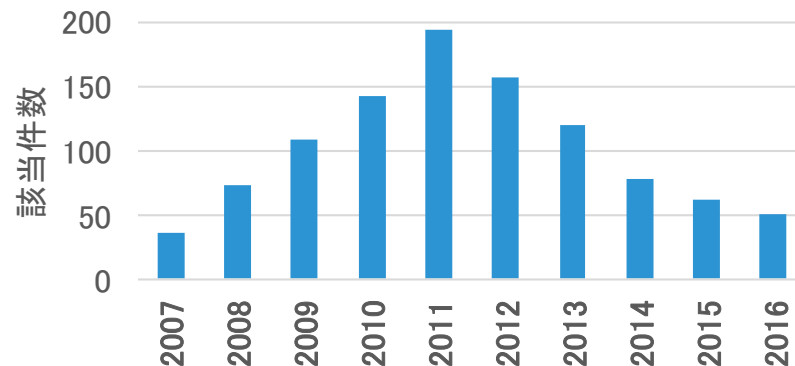
全体の出願件数の推移

全体では緩やかに減少傾向にある



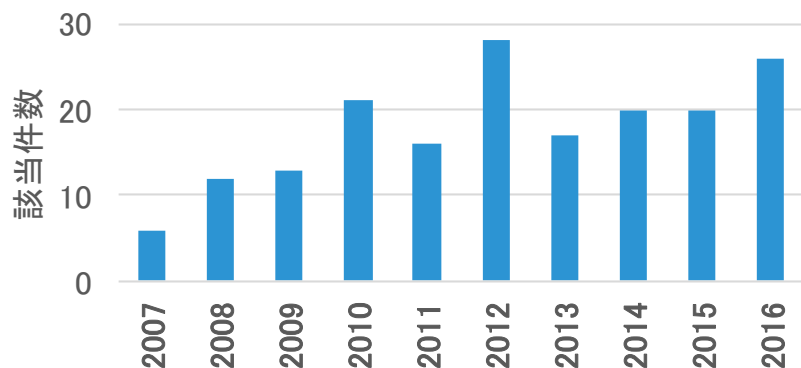
「Ta04.電気自動車への充電、給電装置」の推移

Ta04は2014年以後の該当割合ワースト3位



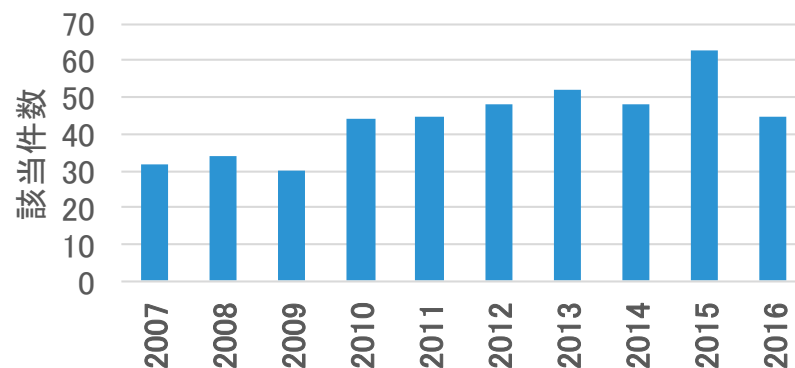
「Ta03.交流・直流の変換」の推移

Ta03は2014年以後の該当割合1位



「Ta07.情報通信、検出判定システム」の推移

Ta07は2014年以後の該当割合2位

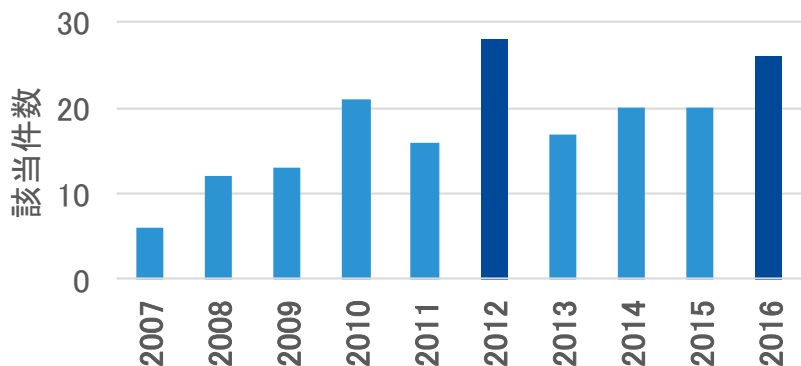


特徴が目立つ特許群のキーワード

該当件数の多いTa03の2012年は電圧変換、Ta03の2016年は電気自動車のバッテリーの電力変換、Ta07の2015年2016年は光通信に関するキーワードが特徴的です

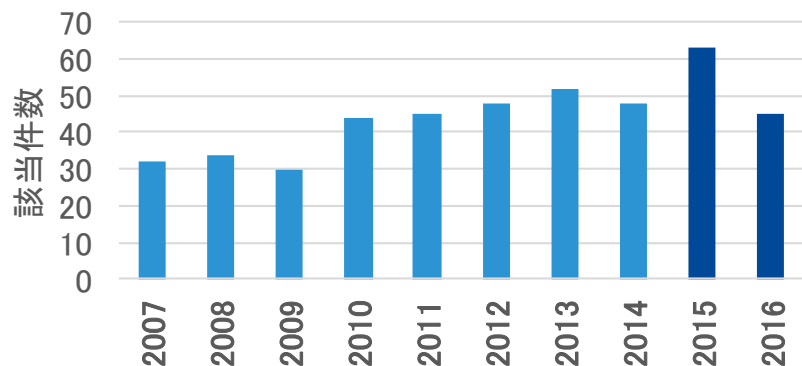
Ta03の該当が多い年の特許のキーワード

Ta03.交流・直流の変換

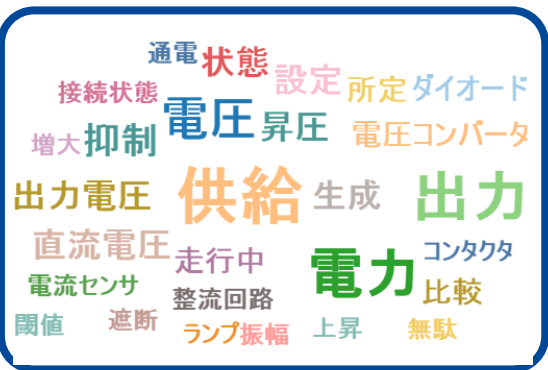


Ta07の該当が多い年の特許のキーワード

Ta07.情報通信、検出判定システム

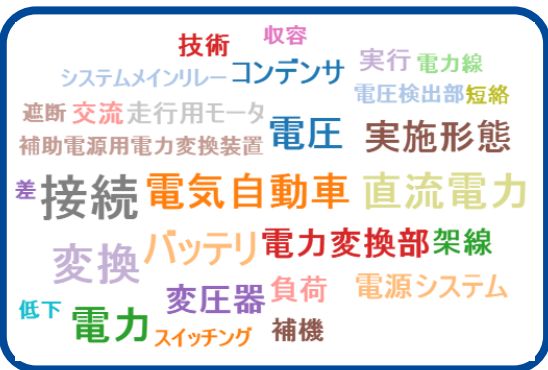


Ta03×2012年の要約キーワード



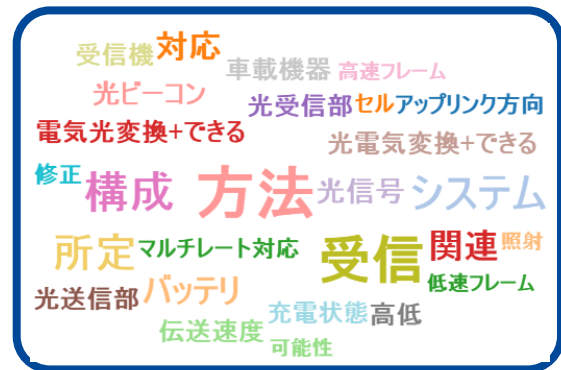
電圧の変換に関するキーワードが多い

Ta03×2016年の要約キーワード



電気自動車やバッテリーの電力変換に関するキーワードが多い

Ta07×2015年2016年の要約キーワード



光信号の送受信のシステムに関するキーワードが多い

PCSAの適用: PIスコアの特徴トピックの抽出

5点以上

該当件数:7,558件 (28.6%)

5点未満

該当件数:18,861件 (71.4%)

分析の
ポイント

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

Tb01.エンジン駆動・動力伝達の制御

作動 回転停止 発電
動力ブレーキ 駆動力 トルク 駆動源
走行 駆動輪 内燃機関車輪 連結 **モータ**
制御装置 ハイブリッド車両 モータジェネレータ
トランスミッション 発電機 **エンジン** クラッチ
制御 駆動 出力軸 開始 入力軸 伝達構成
バッテリー

Tb02.交流・直流の変換

放電 並列モータ 充電 変換 駆動
コンデンサ **電力電圧** **バッテリー** **制御**
給電電気自動車 **電力電圧** **バッテリー** **制御**
制御装置 直流インバータ 検出 電力変換装置
電気負荷 **供給** 交流電力 スイッチ 蓄電装置
電源 直流電力 コンバータ 直列 負荷 制御部
電流

Tb03.電気自動車への充電、給電装置

開始電源
受信 演算 給電装置 充電時 情報 送信 コーダ制御部
取得 充電システム 蓄電池 放電 二次電池 充電ケーブル
電力 走行 **電気自動車** 蓄電装置
充電スタンド 充電+できる 給電 **充電** **バッテリー**
供給 外部電源 制御外部 構成

Tb04.二次電池の製造方法

構成
含有製造方法 バッテリーケース 面 導電性 表面形成
電池特性 **バッテリー** 電池パック 集電体 セパレータ
正極電極 **二次電池** リチウムイオン電池
収容 **負極** 正極活物質 電池モジュール 電解液
被覆 電解質 負極活物質 積層活物質 材料
組電池 方法

Tb05.空調などの冷却・加熱、ガス処理

作動 **方法** 駆動 回転 生成
排気ガス 冷却 内燃機関 熱エネルギー **配置**
排出システム 内燃機関 熱エネルギー **配置**
供給 制御 電気エネルギー 変換 **発電機**
電気機械 **構成** 発電 加熱空気 自動車
電力 エンジンモータ 温度 車室内
バッテリー

Tb06.状態の検出と判定

入力センサ 測定位置 停止
信号 **制御** 生成 比較 ECU送信方法
制御部 **演算** **検出** 値 **判定** 制御装置
情報 受信 異常記憶 構成 電気信号
閾値 駆動 取得変化 電圧 調整

Tb07.形成・配置

接触一対基板一体 下方
方向位置 開口部 一端 **連結** **収容**
固定 **構成** **配置** ハウジング **形成**
回転コネクタ 支持モータ 突出 端部 ケース
外部 軸方向 保持 対向 上方 車体 移動

Tb08.小型化・低コスト化・簡素化・安全性向上

コスト
起因 **構造** 効率 実現 安定 **バッテリー** 操作
信頼性 必要+ない 安全 **確保** 低コスト
小型化 劣化 **電気自動車** 損傷 ハイブリッド車両
自動車 電子機器 **構成** 電気部品 車室内
部品点数 影響 耐久性 スイッチ 確保+できる
振動 安価

Tb09.重力発電の活用による地球温暖化防止

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

Tb10.既存エンジンへの警鐘・タービン発電の提案

飛行機 軽量物発電 運用
安価電気駆動 軽量蒸気速度 横軸+直車 燃費
反転 太陽光加熱器製造 自動車液体酸素圧縮駆動
既存エンジン 宇宙到達費用 **理論最良エンジン**
電気+液体空気+過熱蒸気温熱供給設備D 大学
静翼 **発電量** 永遠 既存蒸気タービン発電
既存 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 全動翼船舶

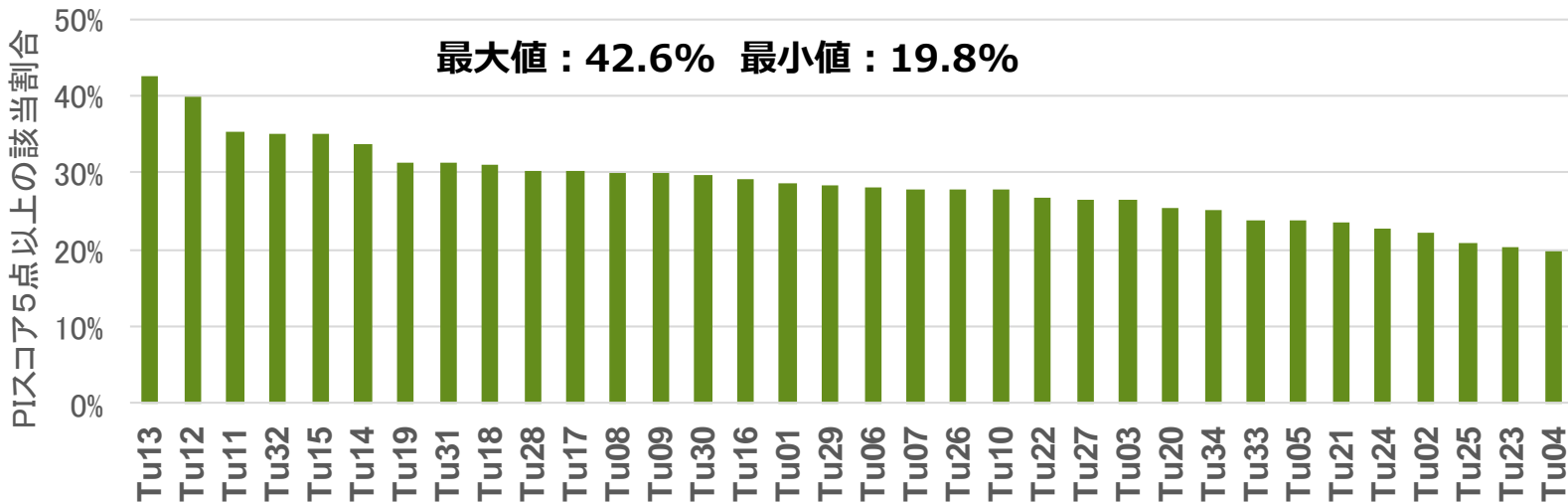
Tb09, Tb10は特定の出願人による重複した要約内容の特許から抽出された

※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）

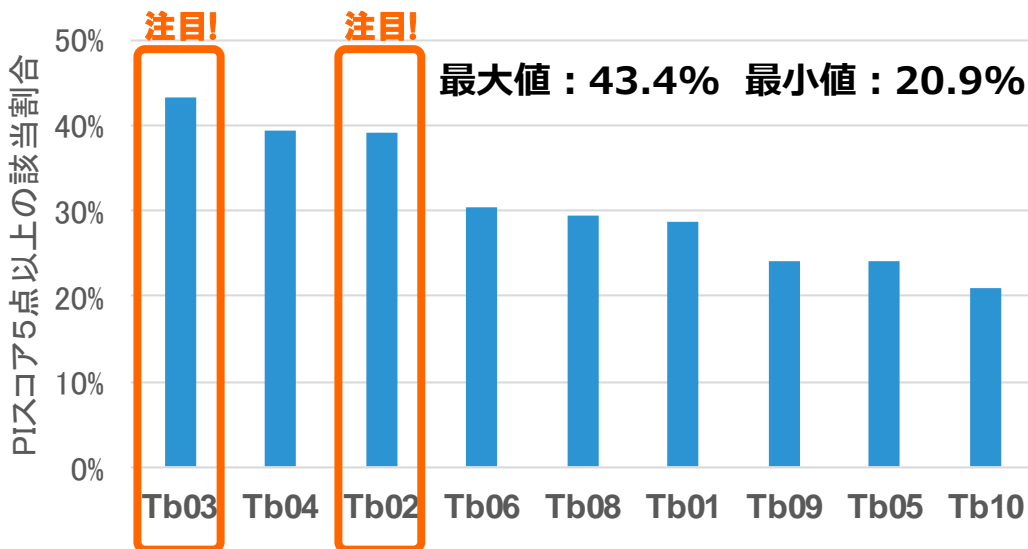
全体のトピックとの属性分布の比較

トピックごとにPIスコア5点以上のデータの該当割合を比較すると、**最大値・最小値はほぼ変わりませんが、全体と比べてPCSAでは平均的な該当割合のトピックの抽出は限定されます**

全体のトピック



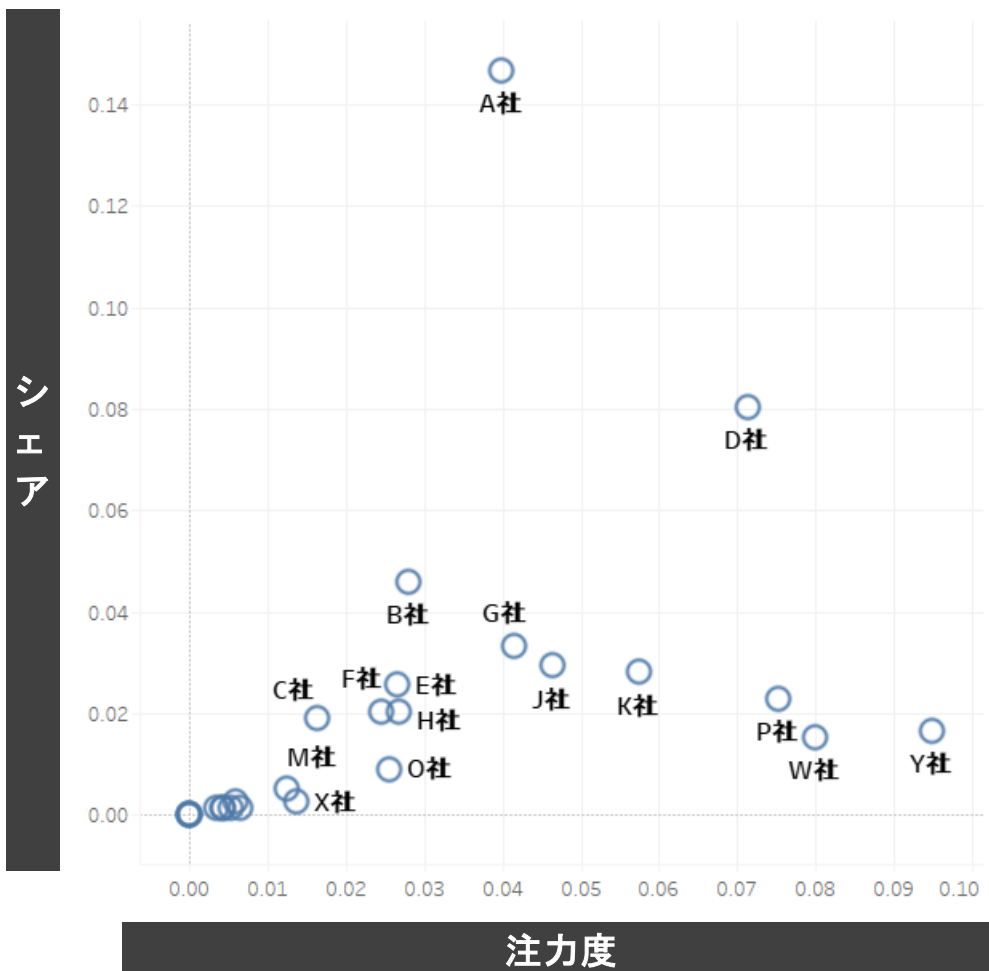
PCSA適用の特徴トピック



- 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
- 全体のトピックもPCSAのトピックも割合の最大値・最小値はほぼ変わらないが、全体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
- Tb07は該当データがなかった
- 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1社と、注力度が高い1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

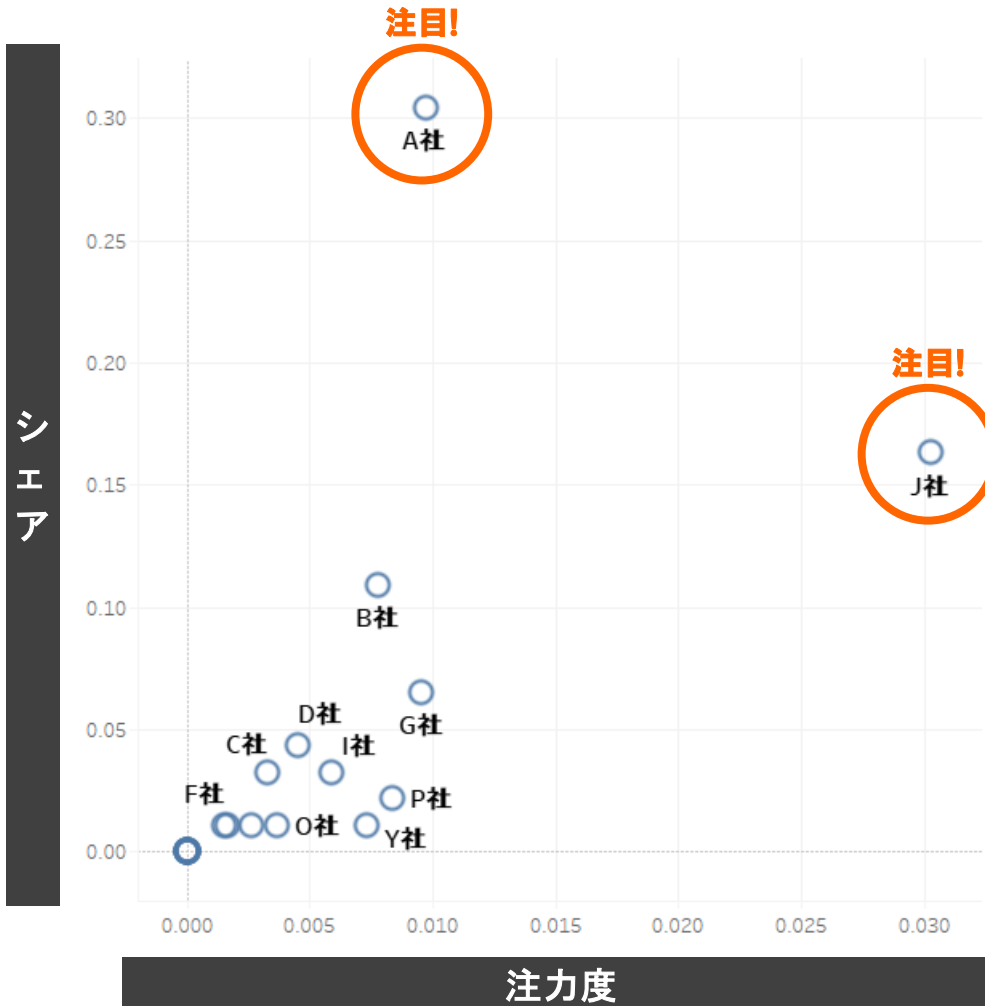
- 最もシェアの高いのはA社で、2番手は少しギャップを開けてD社であるが、このトピックの領域ではこの2社がメインプレーヤーといえる
- D社はA社にシェアで劣るものの、注力度が高く、独自性の強い技術が開発されている可能性がある

注力度とシェア

- **注力度**: $P(\text{トピック}T = 1 \mid \text{出願人}X = 1)$
 - 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**: $P(\text{出願人}X = 1 \mid \text{トピック}T = 1)$
 - 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

交流・直流の電力変換に関するトピック(Tb02)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おおよそ3社で、特に1社の注力度の高さが際立ちます

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

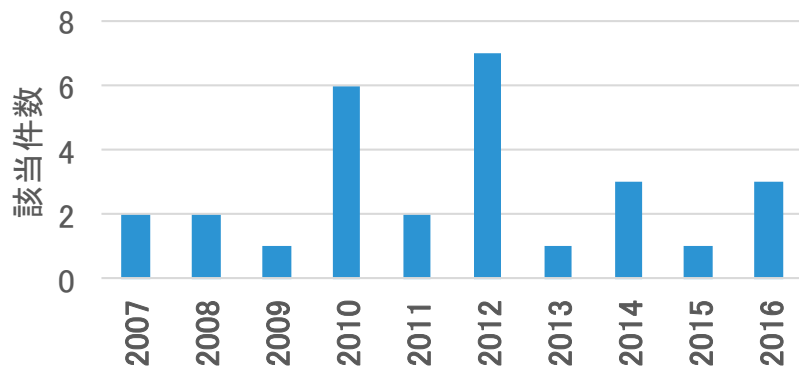
- 最もシェアの高いのはA社で、2番手はギャップを開けてJ社、3番手はさらに少しギャップを開けてB社であるが、このトピックの領域ではこの3社がメインプレーヤーといえる
- 注力度ではJ社がとても高く、他社を大きく突き放しており、独自性の強い技術が開発されている可能性がある

注力度とシェア (再掲)

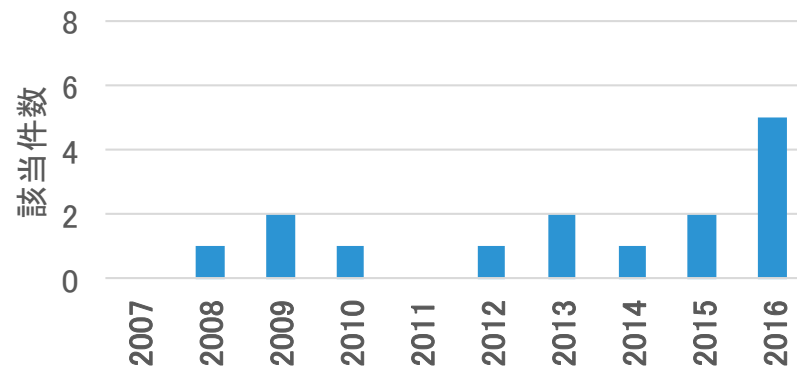
- **注力度**: $P(\text{トピック}T = 1 \mid \text{出願人}X = 1)$
 - 出願人Xの出願特許の中で、どれくらいの割合がそのトピックTに該当するものか、つまり出願人がどれくらいそのトピックに注力しているのかを示している
- **シェア**: $P(\text{出願人}X = 1 \mid \text{トピック}T = 1)$
 - トピックTが該当する特許の中で、どれくらいの割合がその出願人Xの出願によるものか、つまりトピックのなかでどれくらいその出願人が占めているのかを示している

シェアが最も高いA社は電圧変換に関するものが多いですが最近はお願が少なく、注力度の最も高いJ社は直近で出願が増えており鉄道の電力変換に関するものが多いです

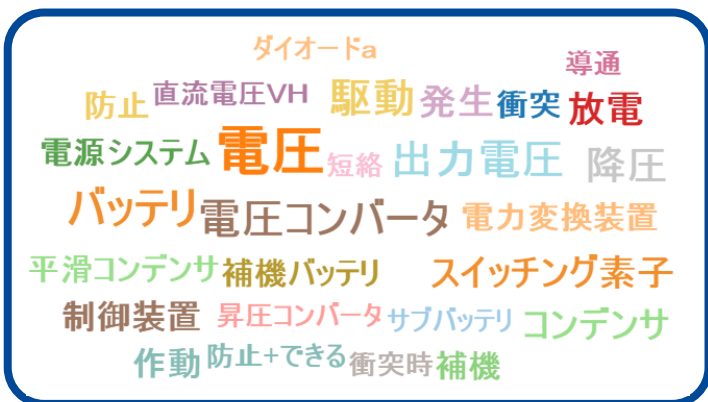
A社(該当28件)の出願推移と要約キーワード



J社(該当15件)の出願推移と要約キーワード

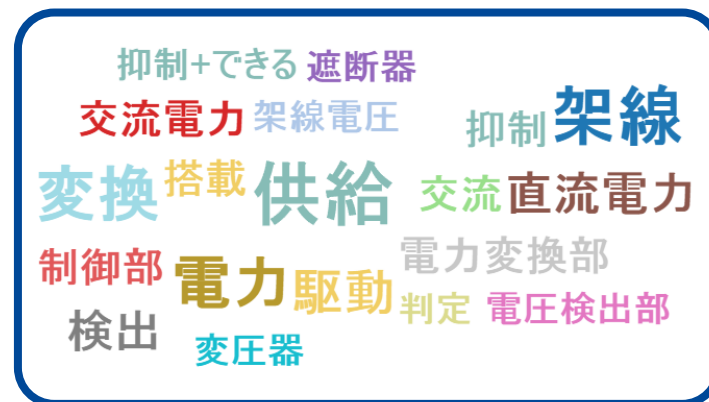


Tb02×A社の要約キーワード



電圧の変換に関するキーワードが多い

Tb02×J社の要約キーワード



架線など鉄道車両の電力変換に関するキーワードが多い

PCSAの適用: 出願人の特徴トピックの抽出

上位5社

該当件数:6,617件 (25.0%)

5社以外

該当件数:19,802件 (71.4%)

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？
出願件数の多い企業とそうでない企業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

Tc01.エンジン駆動・動力伝達の制御

駆動輪動力 運転状態 開始 制御装置
電気式差動部 動力伝達経路
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

Tc02.モータ駆動、ブレーキシステム

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

Tc03.交流・直流の変換

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
リレ-蓄電装置 インバータ バッテリ制御装置
コンデンサ 供給 モータ 駆動 電力変換装置
電源装置 直流電力スイッチ 並列 電流 交流電力
作動

Tc04.電気自動車への充電、給電装置

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

Tc05.二次電池の構成・製造方法

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

Tc06.空気や燃料電池などの温度制御

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

Tc07.状態の検出と判定

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

Tc08.形成・配置

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

Tc09.スイッチの操作性

固定接点 導電ケース 検出 操作体
ストッパ 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 安插 接触

Tc10.小型化・低コスト化・簡素化・信頼性向上

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 安插 損傷 モータ
構成 精度

Tc11.既存エンジンへの警鐘・樹脂組成物の提供

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

Tc12.重力発電の活用による地球温暖化防止

垂直下方 地球温暖化 全係地球温暖化
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 安插 水
工場電化全盛 二酸化炭素排気ゼロ 人類絶滅
海軍 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

Tc13.タービン発電の出力向上・燃費低減

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 安插 電気駆動 最大速度 潜水 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

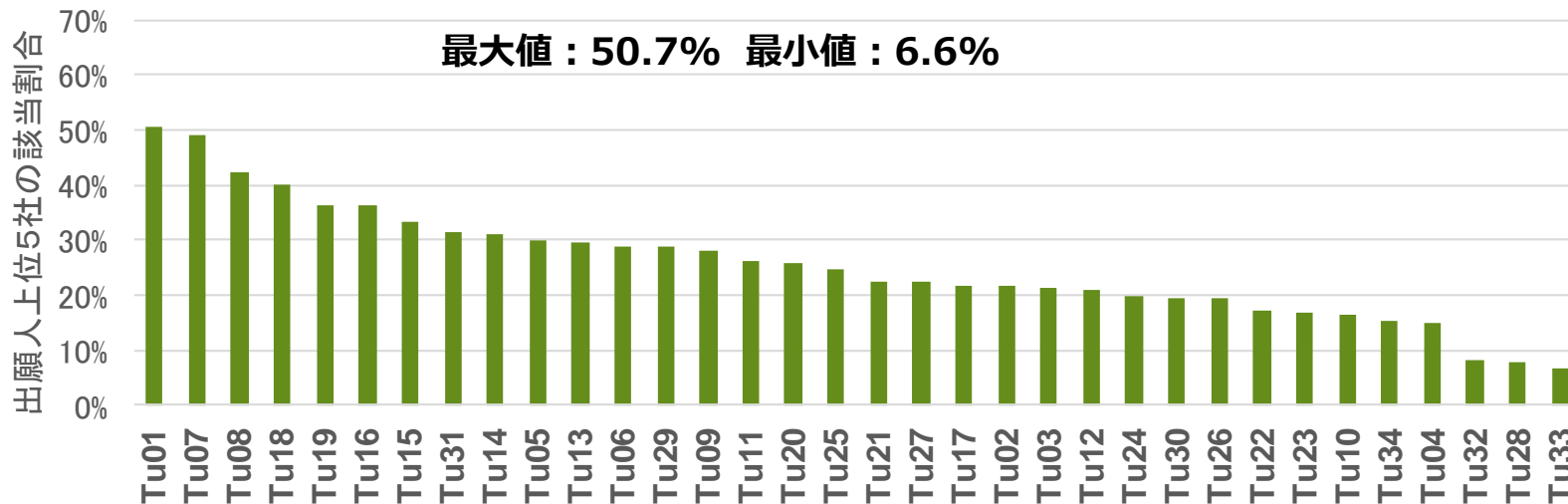
Tc11, Tc12, Tc13は特定の出願人による重複した要約内容の特許から抽出された

※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)

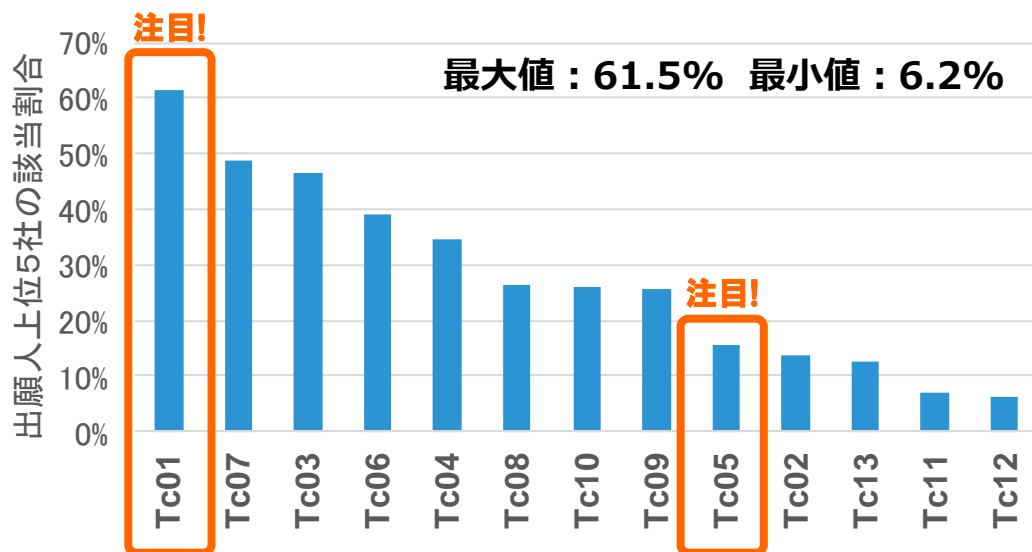
全体のトピックとの属性分布の比較

トピックごとに出願人上位5社データの該当割合を比較すると、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、数も限定的に抽出されています

全体のトピック



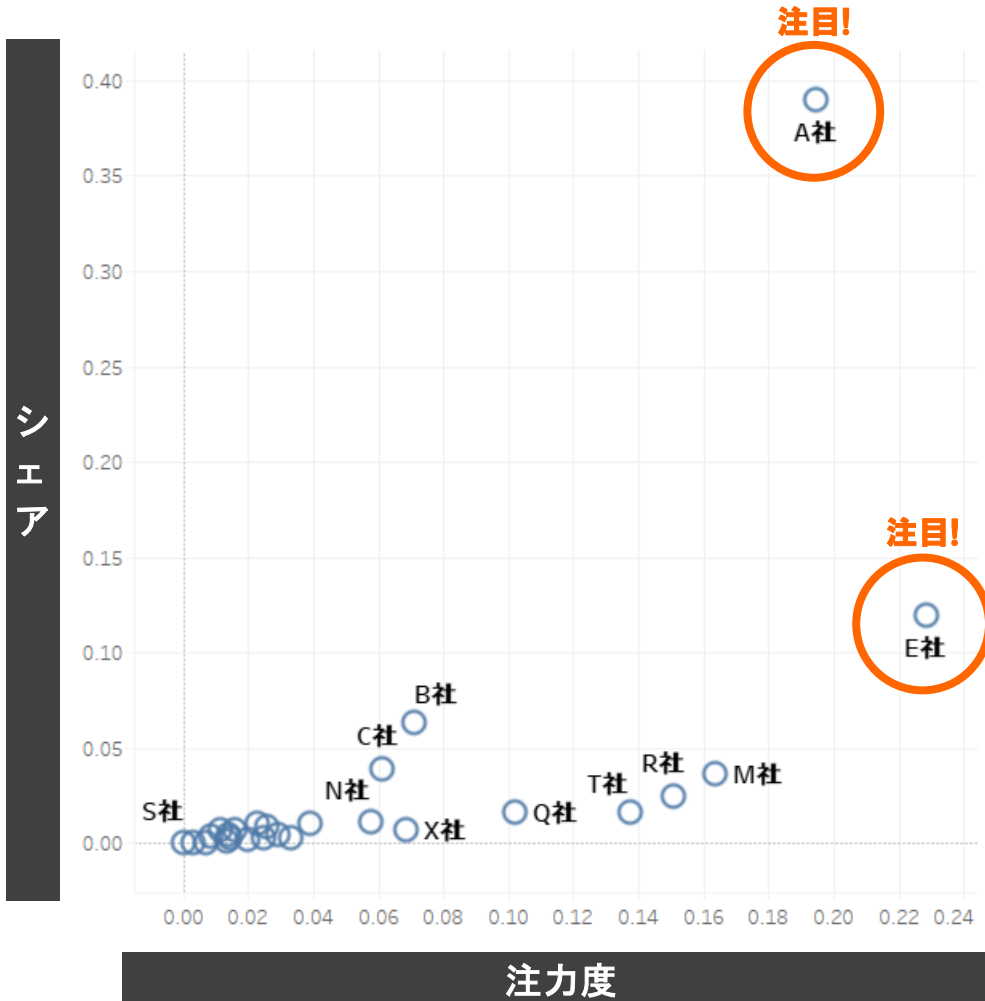
PCSA適用の特徴トピック



- 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
- 全体のトピックもPCSAのトピック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全体よりも限定的に抽出されている
- 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、実際にメインプレーヤーはその5社のうちの2社で、どちらも注力度が高いです

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

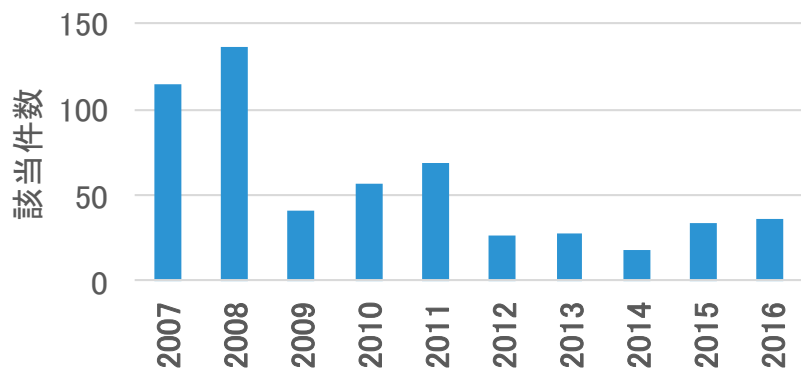
- 最もシェアの高いのはA社で他社を大きく突き放しており、2番手は大きくギャップを開けてE社であるが、E社も比較的高い水準のシェアを保有している
- 注力度も最も高いのはほぼ同じ水準でA社とE社であり、このトピックの領域ではこの2社がメインプレーヤーといえる

注力度とシェア (再掲)

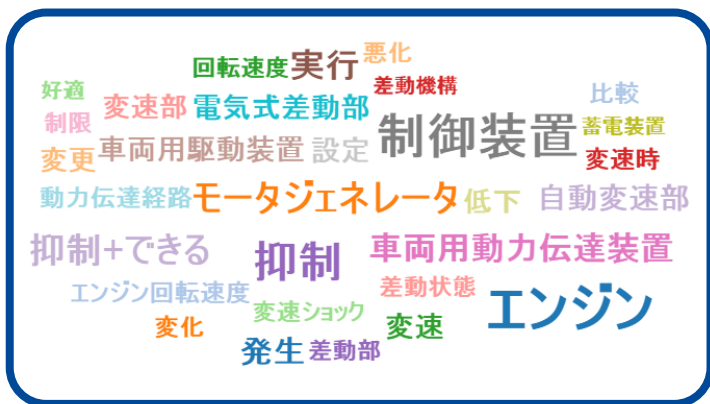
- **注力度**: $P(\text{トピック}T = 1 \mid \text{出願人}X = 1)$
 - 出願人Xの出願特許の中で、どれくらいの割合がそのトピックTに該当するものか、つまり出願人がどれくらいそのトピックに注力しているのかを示している
- **シェア**: $P(\text{出願人}X = 1 \mid \text{トピック}T = 1)$
 - トピックTが該当する特許の中で、どれくらいの割合がその出願人Xの出願によるものか、つまりトピックのなかでどれくらいその出願人が占めているのかを示している

圧倒的シェアのA社は差動や変速に関するものが多いですが出願が多いのは10年前で、シェア2位のE社はハイブリッド車やクラッチに関するものが多いですが最近では出願が減っています

A社(該当560件)の出願推移と要約キーワード

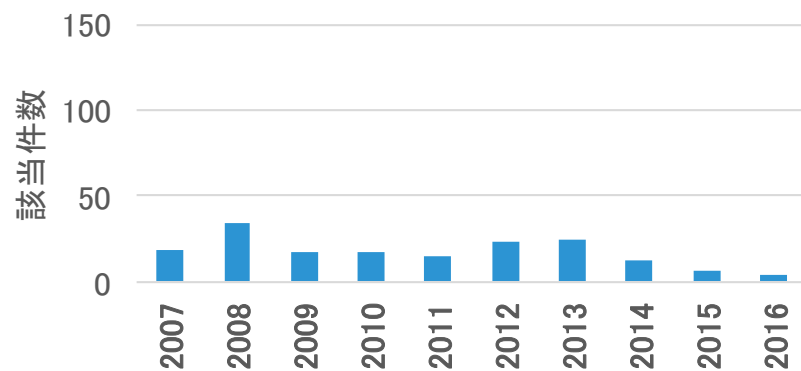


Tc01×A社の要約キーワード

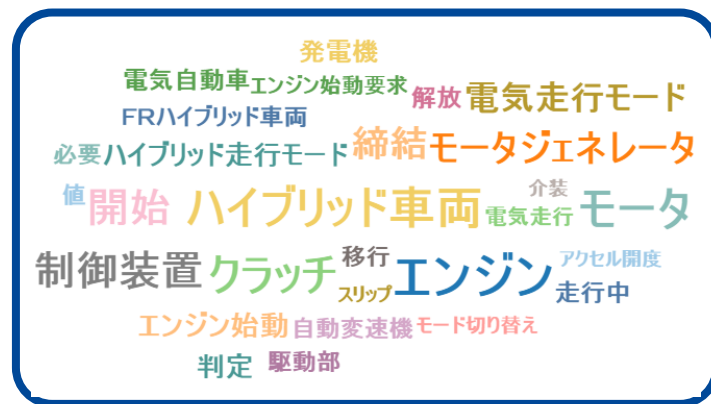


差動や回転速度・変速に関する動力制御のキーワードが多い

E社(該当172件)の出願推移と要約キーワード



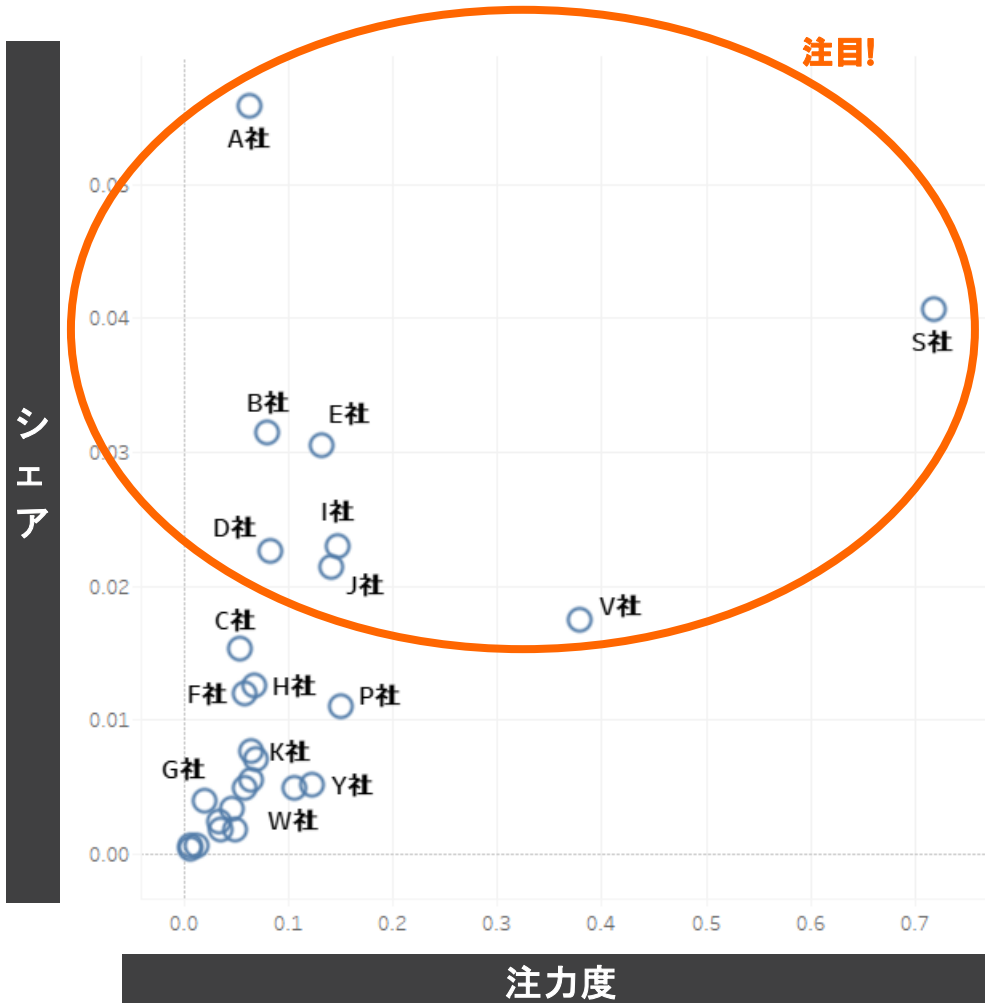
Tc01×E社の要約キーワード



ハイブリッド車両やクラッチに関する動力制御のキーワードが多い

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

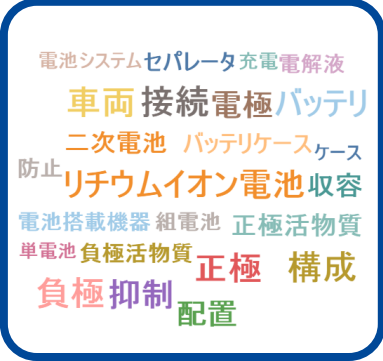
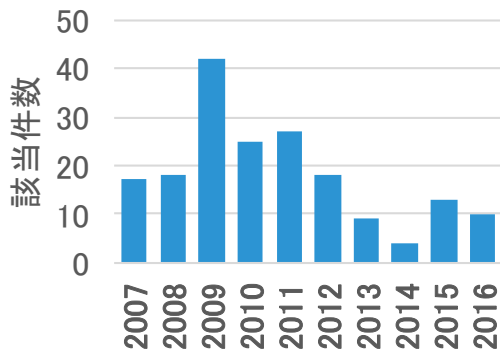
- 最もシェアの高いのはA社で、次いでS社だが、他にも比較的シェアを保有している出願人も複数あり、プレーヤーが多い領域といえる
- 注力度ではS社が最も高く、他社を突き放しており、次いでV社の注力度も高い水準にある
- 巨大企業だけが占有するのではなく、各社が独自の技術を保有して棲み分けがされている可能性がある

注力度とシェア

- **注力度**: $P(\text{トピック}T = 1 \mid \text{出願人}X = 1)$
 - 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**: $P(\text{出願人}X = 1 \mid \text{トピック}T = 1)$
 - 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

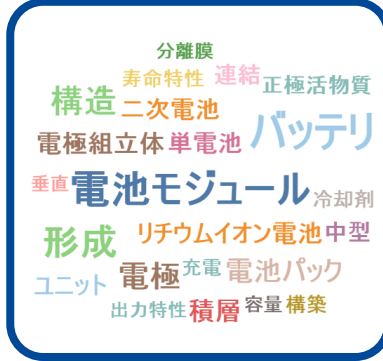
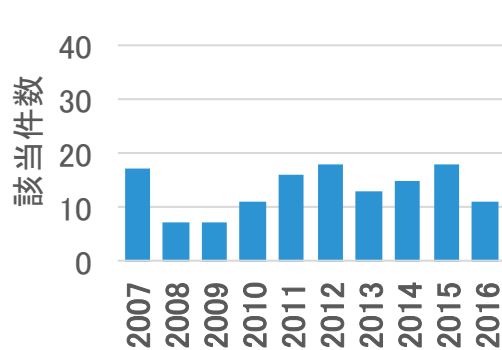
シェア1位のA社は2009年が出願ピークでリチウムイオン電池に関するものが多く、シェア2位注力度1位のS社は年10~20件の安定した出願で電池モジュールに関するものが多いです

A社(該当183件)の出願推移と要約キーワード



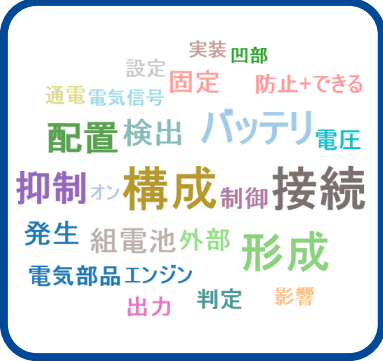
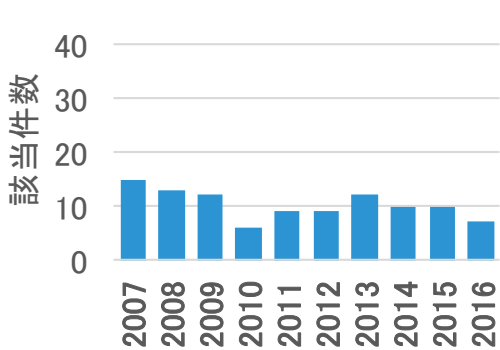
リチウムイオン電池や電極、接続・構成、收容ケースなど二次電池の全般的なキーワードが多い

S社(該当133件)の出願推移と要約キーワード



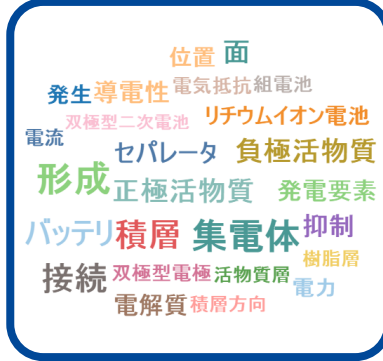
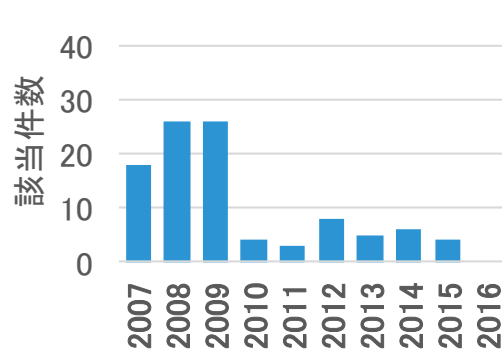
電池モジュールや電池パック、積層された電極組立体などに関するキーワードが多い

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

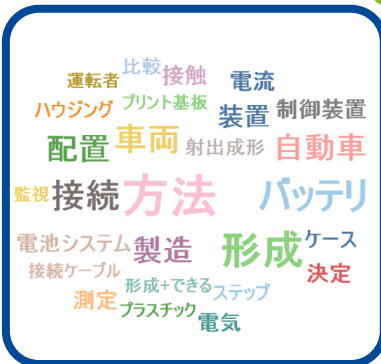
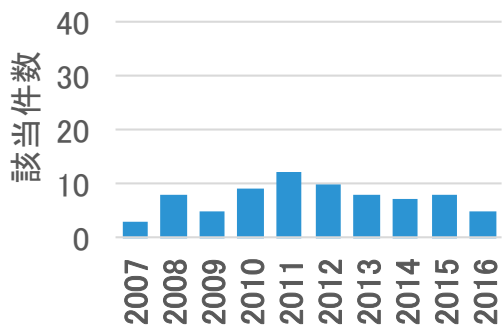
E社(該当100件)の出願推移と要約キーワード



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

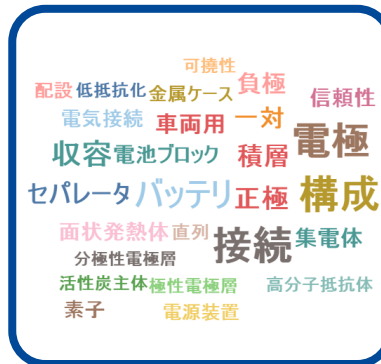
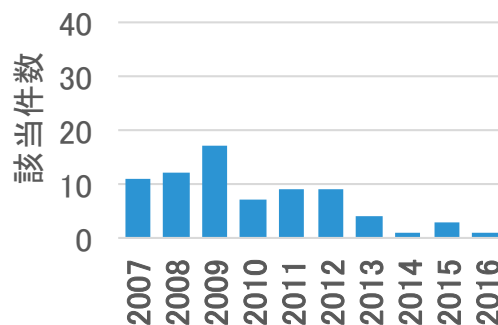
J社はシェア・注力度ともに高い水準ではないですが、直近で出願が急増しており、非水電解質や活物質に関するものが多く、注力度2位のV社は金属材料に関するものが多いです

I社(該当75件)の出願推移と要約キーワード



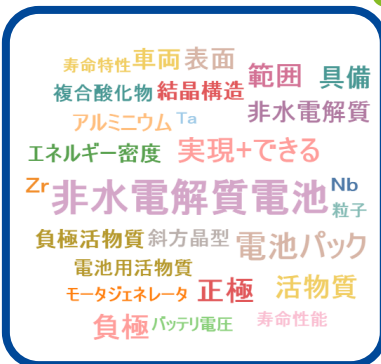
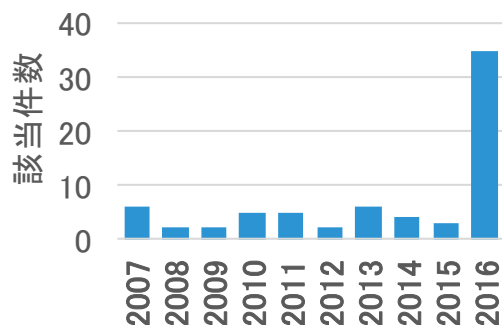
ケース、ハウジングなど筐体に関するキーワードが多い

D社(該当74件)の出願推移と要約キーワード



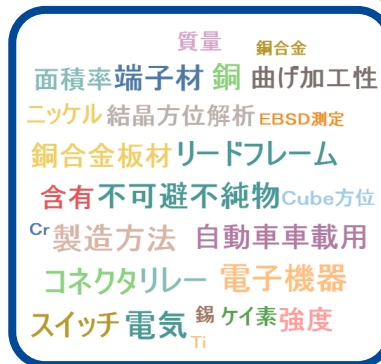
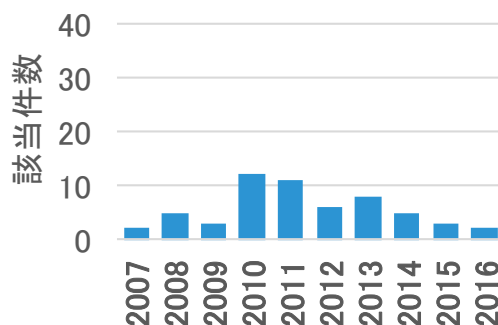
電極、セパレータ、構成などに関するキーワードが多い

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

まとめ

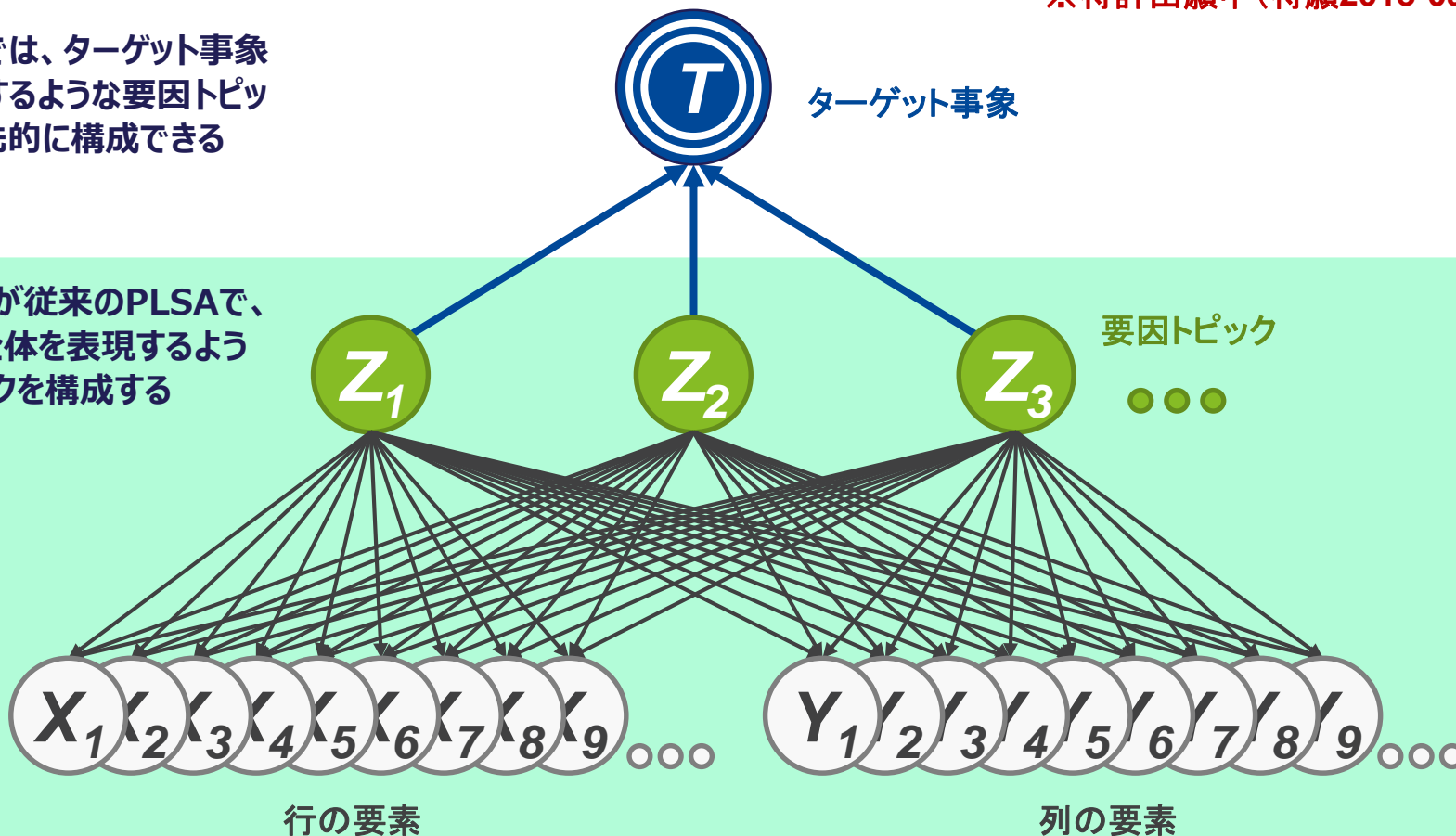
PCSA(確率的因果意味解析)は、ターゲット事象に影響を与える要因トピックを優先的にテキストデータから自動構成でき、よりインサイトに近づく探索軸を獲得することができます

確率的因果意味解析 (PCSA: Probabilistic Causal Semantic Analysis)

※特許出願中(特願2018-090885)

PCSAでは、ターゲット事象に関連するような要因トピックを優先的に構成できる

ここまでが従来のPLSAで、データ全体を表現するようなトピックを構成する



PCSAで特許文書を分析することで、確認したい対象に合わせた特徴要因をテキスト情報から自動構成できるため、それを分析軸に探索することで効果的なインサイト獲得が期待できます

Step1 ターゲット事象の設定

出願年

最近の出願特許の特徴とは？

PIスコア

スコアの高い有用技術の特徴とは？

出願人

出願件数の多い企業で牛耳られている技術の特徴、逆に棲み分けができていない技術の特徴とは？

Step2 差分共起行列の作成とPLSAの適用による特徴トピックの抽出

テキストマイニングによる単語の抽出

| 単語 | 頻度 | 係り受け | 頻度 |
|------|-------|----------|-------|
| 構成 | 4,997 | 電力-供給 | 1,208 |
| 制御 | 4,360 | 否-判定 | 517 |
| 配置 | 3,895 | モーター-駆動 | 460 |
| モーター | 3,486 | バッテリー-充電 | 440 |
| ... | ... | ... | ... |

ターゲット事象の該当有無別に共起行列を2つ作成

| | 電力-供給 | 否-判定 | モーター-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 |
| モーター | 239 | 61 | 494 | 58 |

共起行列U

| | 電力-供給 | 否-判定 | モーター-駆動 | バッテリー-充電 |
|------|-------|------|---------|----------|
| 構成 | 97 | 30 | 27 | 25 |
| 制御 | 209 | 49 | 80 | 66 |
| 配置 | 59 | 1 | 23 | 9 |
| モーター | 192 | 42 | 356 | 46 |

共起行列A

| | 電力-供給 | 否-判定 | モーター-駆動 | バッテリー-充電 |
|------|-------|------|---------|----------|
| 構成 | 21 | 3 | 9 | 8 |
| 制御 | 59 | 24 | 28 | 18 |
| 配置 | 13 | 1 | 6 | 0 |
| モーター | 47 | 19 | 138 | 12 |

共起行列B

| | 電力-供給 | 否-判定 | モーター-駆動 | バッテリー-充電 |
|------|-------|------|---------|----------|
| 構成 | 10.4 | 6.7 | 0.3 | 0.1 |
| 制御 | 7.3 | 4.4 | 0.4 | 0.4 |
| 配置 | 5.1 | 0.4 | 0.4 | 2.6 |
| モーター | 15.2 | 5.4 | 22.7 | 2.9 |

共起行列 A-B

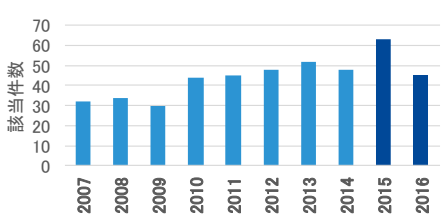
2つの共起行列の差分共起行列にPLSAを適用してトピックを抽出

PLSA

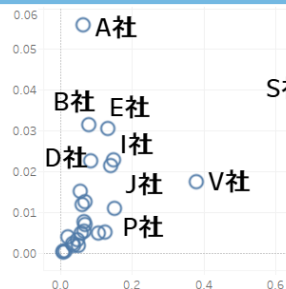
- 交流・直流の変換
 - 交流電力充電電気自動車コンバータ
 - 交流電力充電電気自動車コンバータ
 - 制御 制御部駆動 電力変換装置 蓄電装置
 - コンデンサ 制
- 二次電池の構成
 - 形成リチウムイオン電池
 - 方法 負極二次電池積層電池モジュール
 - 構造コネクタ 正極活性物質 負極活性物質
- 電気自動車への充電
 - 取得充電システム蓄電池 放電二次電池 充電ケーブル
 - 電力 電気自動車 蓄電装置
 - 充電スタンド 充電+できる給電
 - 供給 外部電源 制御 外部 充電 構成

Step3 特徴トピックを軸とした探索的分析

出願件数のトレンドを可視化



出願人のポジショニングを可視化



要約の特徴キーワードを可視化

寿命特性 車両表面 範囲 具備
 複合酸化物 結晶構造 アルミニウム Ta 非水電解質
 エネルギー密度 実現+できる
 Zr 非水電解質電池 Nb 粒子
 負極活物質 斜方晶型 電池パック
 電池用活物質
 モータージェネレータ 正極 活物質
 負極 バッテリ電圧 寿命性能

PCSAの技術は様々な業務のテキストデータに適用でき、効果的なビジネスアクションの検討を支援します



口コミ

評価得点に影響を与える
口コミのトピックとは？



アンケート

顧客満足度に影響を与える
自由記述トピックとは？



コールセンター履歴

解約・退会に影響を与える
問い合わせトピックとは？



特許文書

近年出願が伸びている
技術トピックとは？



営業日報

契約成立に影響を与える
営業活動トピックとは？



有価証券報告書

業績指標に影響を与える
定性的なトピックとは？



エントリーシート

選考通過に影響を与える
自己PRトピックとは？



診療記録

検査指標に影響を与える
定性的なトピックとは？



問題発生レポート

問題の発生に影響を与える
報告記録トピックとは？

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て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

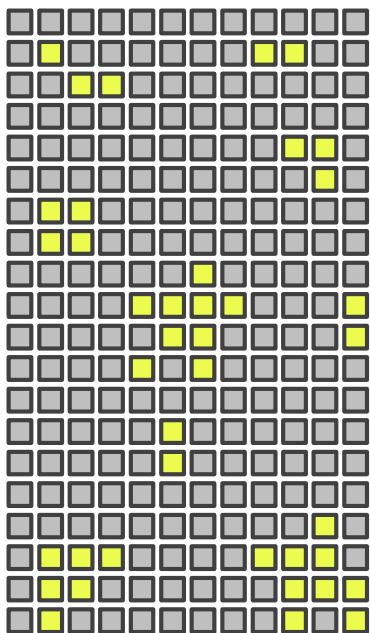
業務課題



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



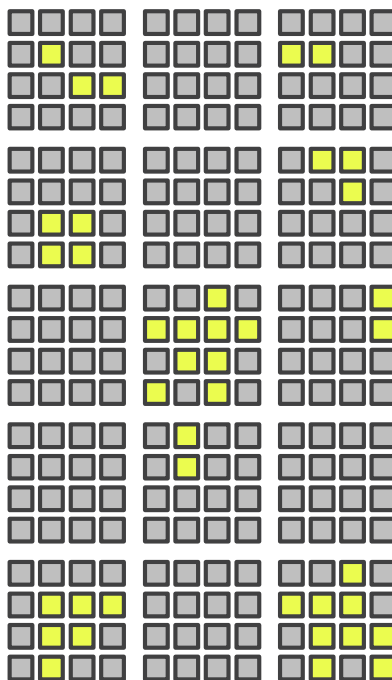
■ 価値あるデータ

マクロ探索アプローチ

特徴の似ているデータの分類

特徴クラスターから個別データの絞り込み

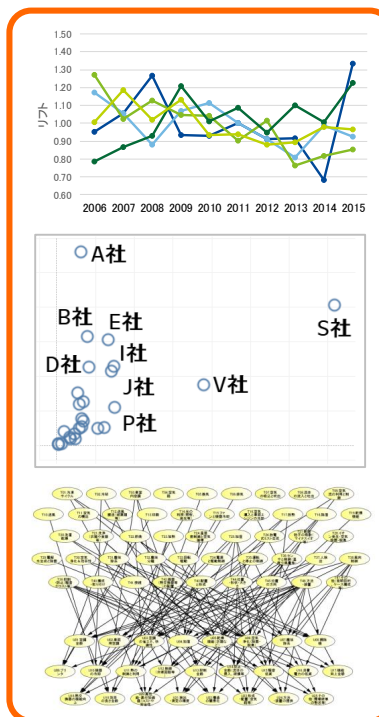
類型化



各クラスターの特徴の可視化

注目すべきクラスターや要因条件の発見

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)

資料に関するお問い合わせは以下までお願いします。

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

