新規性ポテンシャルと単語分散表現モデルに基づく
特許情報からの設計概念生成橘 智也^{*1}, 正田 浩暉^{*2}, Aiza Syamimi Binti Abd Rani^{*4}
野間口 大^{*3}, 岡本 和也^{*5}, 藤田 喜久雄^{*3}Design concept generation from patent information based on
novelty potential and distributional word representation modelTomoya TACHIBANA^{*1}, Koki SHODA^{*2}, Aiza Syamimi Binti Abd Rani^{*4},
Yutaka NOMAGUCHI^{*3}, Kazuya OKAMOTO^{*5} and Kikuo FUJITA^{*3}^{*1,*2,*3} Department of Mechanical Engineering, Graduate School of Engineering, Osaka University
2-1 Yamadaoka, Suita-shi, Osaka 565-0871, Japan^{*4,*5} Graduate School of Innovation and Technology Management, Yamaguchi University
2-16-1 Tokiwadai, Ube-shi, Yamaguchi 755-8611, Japan

Received: 13 July 2021; Revised: 18 October 2021; Accepted: 20 December 2021

Abstract

Generating novel design concepts in the conceptual design phase is a cornerstone for producing innovative products. This paper proposes a method to support the generation of new design concepts from patent documents. The proposed method is based on the theory of novelty potential that the combination of abstract concepts leads to the generation of the novel design concept and that the distance between the combined abstract concepts correlates to the possibility of generating it. This research adopts the distributed word representation model of word2vec to extract abstract concepts from patent documents and to measure the concept distance. The two matrices are introduced to visualize the relationships of the extracted abstract concepts; the novelty potential matrix which shows the value of the novelty potential between the abstract concepts, and the void matrix which shows the numbers of existing patents of the abstract concepts. They help a designer to identify the combinations of abstract concepts that are high novelty potential and no existing patent. This paper demonstrates a case study of generating novel ideas of sports business with blockchain using the proposed method. The results show its uniqueness and effectiveness.

Keywords : Patent information, Design concept generation, Creativity, Design theory, Word2vec

1. 緒 言

特許文書データベースは多くの技術的知識が言語情報として集約された大規模データであり、データサイエンス技術を用いて、そこから新たな技術開発の機会を発見することを目的とする研究が様々に行われている (Luo et al., 2018). 従来、特許文書に含まれる特許番号や出願日、被引用数などの特許間で意味や形式が統一されている構造化データを用いた分析が主流であったが、近年では自然言語処理の分野の発達により発明の内容にまで踏み込んだ分析を大規模に行うことが可能となっている。一方で設計工学の分野では、人間の設計行為のモデルに基づいた設計概念生成支援についての研究が行われている。例えば、著者らのグループでは一般設計学における概念操作モデルに立脚し、抽象概念の抽出と操作を基本とする設計概念生成手法である Externalization - Classification - Finding voids (ECF) サイクルを提案している (野間口他, 2017). さらに、より距離の遠い抽象概念を組み合わせることにより新規性の高い設計概念が得られる (Taura and Nagai, 2013) という考えをもとに、生成さ

No.21-00222 [DOI:10.1299/transjsme.21-00222], J-STAGE Advance Publication date : 7 January, 2022

本論文は、第30回設計工学・システム部門講演会 講演論文集(2020), No.3201の掲載内容に基づいた論文である。

^{*1} 学生員, 大阪大学大学院工学研究科機械工学専攻 (〒565-0871 大阪府吹田市山田丘 2-1)^{*2} 大阪大学大学院工学研究科機械工学専攻^{*3} 正員, フェロー, 大阪大学大学院工学研究科機械工学専攻^{*4} 学生員, 山口大学大学院技術経営研究科 (〒755-8611 山口県宇部市常盤台 2-16-1)^{*5} 山口大学大学院技術経営研究科

E-mail of corresponding author: tachibana@syd.mech.eng.osaka-u.ac.jp

れる概念の新しさの可能性を示す新規性ポテンシャル (Novelty Potential, NP) の指標, および, その計量化を自然言語処理手法の一つである単語分散表現モデルを用いて行う手法を提案している (野間口他, 2020). なお, ここでの抽象概念 (abstract concept) は一定の範囲の設計解を特徴付けている共通的な性質や関係, 状態などを抜き出して指示する言明を意味し, 個別の設計解はそれらの組合せにより特定することができるとする. それらの考え方のもと, 特許情報に対して ECF サイクルを適用し, その中で NP の計量を行えば, 新しいアイデアが生まれる可能性の高い技術分野を機械的に列挙し, 設計概念を効率よく体系立てて生成することができるようになることが期待できる.

本研究の目的は, 上記の可能性を受け, 特許情報から設計概念を生成する手法を新たに構築することである. 提案手法においては初めに特許情報を収集し, 単語分散表現モデルを用いたアプローチにより, 抽象概念の分類・抽出を行う. つぎに, 新しい抽象概念の組み合わせを見つけるためのツールとして NP マトリクスと Void マトリクスを提案し, それらに基づく設計概念生成手法を提示する. 本手法の有効性を検証するため, ブロックチェーンを用いたスポーツビジネス領域の特許を題材として取り上げる. 特許データベースから欧米特許分類 Cooperative Patent Classification (CPC) を活用して特許情報を収集し (Shibata et al., 2018), 本手法を用いて新しいビジネス領域の設計概念の生成を行う. 生成されたビジネス領域に関するキーワードをもとに, 実際の特許出願数動向と比較することで検証結果を考察し, 提案手法の有効性を検証する. 最後に, 今後の展望を述べる.

2. 特許情報活用と設計概念生成についての従来研究の課題と本研究のアプローチ

2・1 特許情報を活用した技術予測支援とその課題

特許情報が技術開発・製品開発において活用される理由は, 以下の4つの点で説明することができる.

- ・ 特許は, 独創性・技術の実現性・商業的価値の明確な基準を満たすことを前提としているため, 技術的属性と市場的属性の両方を有している (Kuznets, 1962).
- ・ 様々な特許データベースの登場により, 大規模に体系的なデータ検索が可能になっている.
- ・ 特許は, 国際特許分類(IPC) 等に基づいて体系的に分類されているため, 特定の技術や製品の分析が容易である.
- ・ 特許情報を分析することで, 競合他社の技術戦略および製品戦略に関する情報が得られ, 競合他社の革新性等を評価することができる (Yang, 2013).

現在, 技術の開発や動向調査, 新規ビジネスの考案などの様々な目的において, 特許情報を活用する研究が多く行われている. 特に, ある企業が自社の持つ技術を活かし, 参入できる可能性のある事業領域は技術機会とも称され, その効率的な探索の手法が模索されている. 例えば, 中小企業の各固有技術に基づいた技術機会を特定するために, 特許情報を用いた分析アプローチ (Lee et al., 2014) や特許データベースと商標データベースに基づいた技術機会の特定を行う (Lee et al., 2020) アプローチなどが研究されている.

特許情報の分析に用いられる項目は2つのカテゴリーに分類される. 1つは, 特許番号, 出願日, 発行日, 譲受人など, 特許間で形式が統一されている構造化された項目であり, もう1つは, 発明の説明内容を表すテキスト情報や図面などの非構造化項目である. 構造化された項目を用いた分析には, 特許の被引用数を用いた分析 (Albert et al., 1991) や分類コードを用いた分析 (Leydesdorff et al., 2014) などがあるが, 特許の発明内容に関する情報を利用していないために説明力には限界がある. そこで近年では言語処理分野の技術を用いて, 特許の発明内容に注目した分析も行われている. 特許文書には図面も含まれているものの, 基本となる内容は言語で説明されており, 自然言語処理に適している. 特許文書の分析では同義語の解釈や複合語の判定, 専門用語の解釈の難解さなどの特許文章特有の課題が存在している. それらの課題に対して, 例えば, Lee らは特許文章中における単語の出現回数を利用し, 二次元マップに落とし込むことにより特許の発明内容の可視化を行っている (Lee et al., 2009). また, 特許の分類の自動化など, より高度な分析を行うために機械学習を用い, 自然言語処理技術の特許に適用する研究が行われている (Abdelgawad et al., 2019).

上記の特許情報の活用についての研究は主に技術機会発見のためのトレンド予測を目的としており, 基本的には既存の特許に含まれる技術のトレンド分析を行うものといえる. 既存の概念を踏まえて, いまだ存在していない新しい概念を生成するためには, 設計概念生成の理論を踏まえた上で, 特許情報の特徴を考慮してその活用を

行う必要がある。特許情報をアイデア生成に用いた研究には、特許の被引用数や分類コードなどの構造化されたデータを用いて作成したパテントマップを利用したのがあり、特許情報をアイデア生成の手段として用いる有用性が認められている (Luo et al., 2018)。一方で、パテントマップの作成方法や、そのレイアウトによりアイデア発想の結果に影響を与えるなどの可能性が指摘されている。また、前段落で述べた技術機会発見の研究と同様、特許の構造化されていないテキスト情報は用いられておらず、そこに含まれる具体的な内容を踏まえたアイデア生成に課題が残っている。

2・2 新規性ポテンシャルと単語分散表現モデルに基づく設計概念生成支援手法

2・2・1 ECF サイクル

一般設計学では、設計のプロセスを概念操作と捉え、集合論を用いて数学的に議論することにより形式化している (吉川, 2020)。その中では設計者が新しい概念を生み出すきっかけとなるのは抽象概念の新しい組み合わせであることが示されている。この抽象概念を組み合わせたときに、実体概念が存在しない空白の組み合わせを Void と呼び、Void の発見が新たな設計概念を生成するきっかけとなる。この一般設計学の Void の考え方をもとにした具体的な発想プロセスのモデルとして ECF サイクルが提唱されている (野間口他, 2017)。ECF サイクルは外在化 (Externalization)、分類操作 (Classification)、Void 発見 (Finding voids) の 3 つの段階から構成されており、設計概念の発想において重要とされるアイデアの外在化と内省的思考の相互作用を用いた発想支援手法である。本手法は有用な発想支援手法の枠組みである一方で、各段階において使用される手法については課題が残されている。例えば、サイクルの前提となる最初の外在化段階でのアイデアの与え方は設計者に委ねられる。概念分類段階では、概念の階層的な分類に用いる機能オントロジーの整備が課題となる。Void 発見段階では 2 つの抽象概念の組み合わせに二元表による視覚的な方法を提案しているものの、組み合わせる抽象概念の選択は設計者の判断に依存する。

2・2・2 新規性ポテンシャルの計量化

ECF サイクルでの Void 発見段階の課題に対して、野間口らによる先行研究 (野間口他, 2020) では、他分野の知識からの類推により設計概念の発想を行う場合、より距離の遠い分野の知識を使うことで、新規性の高い設計概念が得られる (Taura and Nagai, 2013) ことを利用し、類推により新しい設計概念が生み出される可能性の大きさを定量的に測る指標として、新規性ポテンシャル (Novelty Potential, NP) の考え方を提唱した。先行研究では NP を「2 つの機能概念の両方を持つ実体概念 (自然物と人工物の両方を含む) を設計者が思いつく難しさの度合」と定義し、機能概念を表す言語表現に対して単語分散表現モデルを用いることで新規性ポテンシャルの計量化を行っている。単語分散表現モデルは単語をベクトルで表現する手法であり、word2vec (Mikolov et al., 2013) を中心に自然言語処理の分野で注目されている。テキストデータの中で近接して用いられる単語は近い意味を持っているとの考えのもとで、ベクトル空間でも近接するように表現される。これにより単語間の距離を定量的に扱うことが可能となる。先行研究による評価実験 (野間口他, 2017) では、人間の評価による NP の度合いと計量化を行った結果の間に相関が見られることが報告されている。この計量化手法を前項で述べた ECF サイクルにおける Void 発見段階で用いることで、組み合わせる 2 つの抽象概念を選択する際の指針となると考えられる。

2・3 本研究のアプローチ

2・1 節で述べたように特許は技術的知識情報の集合であり、そのすべてが言語化された形で記述されていることが特徴である。本研究では、著者らのグループが提案している ECF サイクルの中で特許情報を活用し、技術的知識に基づいた設計概念生成の方法を確立することを目指す。具体的には、2.2 節で述べた課題について、以下のアプローチで取り組む。その概要を図 1 に示す。まず、初期外在化段階において、一つの特許を一つの実体概念とみなし、特許情報データベースから収集した技術的知識を最初の実体概念として利用する。分類段階においては、単語分散表現モデルを用いて、収集した特許情報を特徴づける抽象概念を抽出し、それによって特許情報を分類する。Void 発見段階においては、抽象概念の間の NP を考え、先行研究で提案した NP 計量化手法を用いて、NP の高い抽象概念の組を網羅的に調べる。設計者は、それらに対して新たなアイデアを考えることができる。

Fig. 1 Concept operation from patent data in ECF cycle.

Fig. 2 Design concept generation process.

3. 特許情報を用いた概念操作による設計概念生成手法

本手法における概念生成の過程について述べる。本手法の全体の流れは図2のように構成されており、(1)から(7)の全7つのステップにより構成されている。以下の各節では各ステップの詳細について述べる。

3・1 ステップ (1) : 特許情報の収集

ステップ (1) では、分析を行う技術領域に関する特許の要旨の文章を集める。本研究では特許に付随されたCPCを扱う。ここで収集する文章はすべて英語で記述されている。収集方法としてはキーワード検索による収

集を行う。本研究では、発明の権利よりも技術的内容に焦点を当てているため、特許請求の範囲ではなく当該技術の骨子が明示された「要旨」部分を特許情報として用いることとする。

3・2 ステップ (2) : 特許情報の前処理

ステップ (2) では、特許文章内に含まれる概念の分類精度を向上させるため、以下の手順で前処理を行う。

- 手順 1. 文章内に含まれる分析に不要な数字・符号を取り除く (例: [,], #, &, %, <, >, 等)
- 手順 2. 複合語を分割する. (例: “blockchainbased” → “blockchain based”)
- 手順 3. 文中に含まれるストップワードに該当する単語を除去する.
- 手順 4. 文章をピリオド (.)および、セミコロン (;) で分割し 1 文とする.

上記の手順 3 では Python の自然言語処理用ライブラリである Natural Language Toolkit (NLTK) (Bird, 2006) に用いられている英語用ストップワードを使用する. 手順 4 において、本研究では、特許の要旨の 1 文が 1 つの抽象概念を表しているとの前提に立ち、文章をピリオド・セミコロン単位で区切る.

3・3 ステップ (3) : 抽象概念の階層構造の構築

ステップ (3) では、抽象概念の階層構造の構築を以下の手順で行う。

- 手順 1. 単語ベクトルを正規化し、1 文における単語ベクトルの平均を取り、それを文ベクトルとする.
- 手順 2. 文ベクトル間のコサイン距離を計算する.
- 手順 3. デンドログラムによる階層構造の構築

上記の手順 1 では word2vec による単語分散表現モデルを用い、文ベクトルを定義する. word2vec は、自然言語処理手法の一つであり、大量の文書データから機械学習を用いて単語をベクトル表現する手法である (Mikolov et al., 2013). ここで学習モデルに含まれる単語の出現頻度とそのノルムについては単語によって差が存在する (Schakel and Wilson, 2015) ことに留意し、式 (1) による単語ベクトルの正規化を行う. さらに 1 文に含まれるすべての単語ベクトルの平均をとったものを文ベクトルとし、式 (2) により定義する. ここで式 (1) における w は word2vec により表された単語分散表現モデルである. 式 (2) における n は一文に含まれる単語数を表す.

$$w_{\text{norm}} = \frac{w}{|w|} \quad (1)$$

$$u = \frac{1}{n} \sum_{i=0}^n w_{\text{norm } i} \quad (2)$$

手順 2 では式 (3) によって定義されるコサイン距離を用いて、すべての文ベクトル間の距離を計測し、各概念間の距離として扱う.

$$\text{Cos distance} = 1 - \frac{uv}{|u||v|} \quad (3)$$

ここで u, v は本節の手順 1 で定義した文ベクトルを表す. 加えて v はコサイン距離により文ベクトル間の距離を計るための、 u に対するもう一方の文ベクトルである. コサイン距離は文章のクラスタリングにおいて多く用いられる計算手法である. コサイン距離は 2 つの文の内容の傾向の類似度を表し、その値が小さいほど距離が近く、2 つの文同士が似た内容の傾向であることを意味する.

手順 3 では概念間の距離の近い抽象概念同士のグループが、より上位の抽象概念に対応すると考え、概念の階層木を構築する。概念のグルーピング手法として、ward 法を用いたデンドログラムを用いる。デンドログラムを描画するメリットとして、グルーピングの最終結果までの過程を視覚的に追うことが可能であること、距離情報を用いたグルーピングが可能であること、ステップ (4) で行う抽象度の設定において最終的にいくつかの概念集合に分けるかということが視覚的に検討しやすいことが挙げられる。

3・4 ステップ (4) : 抽象度の設定

ステップ (4) では、デンドログラムとして得られた概念階層構造に対して、閾値を設定して抽象概念のクラスタを形成する。一般的にデンドログラムを用いた階層的クラスタリングにおける適切な閾値、及びクラスタ数の決定は、データの性質を理解した上で検討する必要がある。定量的な閾値の決定法としてはシルエット法 (Rousseeuw, 1987) 等があり、それらはある程度客観的な指標を与えるものである。一方で、本研究のような文章により構成される概念のクラスタリングでは、明確なクラスタの境界があるわけではないため、人間が判断するのが適切である。本研究では、1 つのクラスタが 1 つの抽象概念に対応すると考え、クラスタに含まれる内容を人が判断し、閾値を決定する。その際、閾値を小さくするほど形成される概念は独立する傾向を示し、その内容はより具体的なものとなる。逆に閾値を大きくするほど概念は結合する傾向を示し、その内容はより抽象的なものとなる。次に、形成されたクラスタに対応する抽象概念を付与する。本研究ではこの作業を手動で行うこととする。具体的には、分類されたクラスタにおける文を人間が通読し、文から読み取った情報を総称する抽象概念を考慮し命名を行う。

3・5 ステップ (5) : 新規性ポテンシャルマトリクスの作成

ステップ (5) では、ステップ (4) で設定した抽象概念間の NP を計量する。本研究ではデンドログラムによる抽象概念間の結合距離を概念距離、すなわち NP とする。計量化した NP の値を、縦横に抽象概念を配したマトリクスの各セルに配置し、さらに NP の値の比較を視覚的に行えるようにするため、ヒートマップを用いた描画を行う。これを NP マトリクスと呼ぶ。本定義により、NP マトリクスは対角要素を持たない対称行列となる。

3・6 ステップ (6) : Void マトリクスの作成

ステップ (6) では、各特許要旨がどの抽象概念を含むかなどの情報をもとに Void マトリクスを作成する。本手法では特許要旨を構成する文、すなわちデンドログラムの最下層の抽象概念の上位の抽象概念が、その特許の特徴を表す抽象概念であるとみなす。一般に、特許と抽象概念は多対多の関係になる。縦横に抽象概念を配したマトリクスを作成し、各セルに対応する特許の数を示したものを Void マトリクスと呼ぶ。1 つの特許に A という抽象概念と B という抽象概念が対応する場合、A と B の概念をもつ特許が 1 つ存在するとして、A と B の交差するセルの数値に 1 加えるなどの操作を行う。以上の操作を、対象とする特許要旨すべてに適用することにより Void マトリクスが完成する。Void マトリクスについても定義より対称行列となる。

3・7 ステップ (7) : 新たな設計概念の生成

3・5 節で作成した NP マトリクスでは概念間の距離が遠い組み合わせを選ぶことで新規性の高い設計概念を生成できることが期待できる。3・6 節で作成した Void マトリクスでは存在している特許が少ない概念の組み合わせを選ぶことで、今まで存在しなかった設計概念を生成できることが期待できる。設計者は、これらの両方のマトリクスを用いて、「新規性ポテンシャルが高い」かつ「今まで存在しなかった」設計概念の組み合わせを網羅的に特定することができる。一般に、そのような設計概念の組み合わせは多数存在する。本研究では、その中から設計者が、具体的な着想ができる組み合わせを選択し、対応する実体概念を生成する。本研究では、この作業は直接の支援の対象とせず、設計者の判断の下で行うこととする。

Fig. 3 Classification of concepts by dendrogram. The horizontal items are abstract concepts. The vertical line shows the cosine distance between the abstract concepts. 16 abstract concepts were extracted under a threshold value of 0.87.

4. 本手法の実施例

4・1 概要

提案手法の検証のため、スポーツビジネスに関わる研究においてブロックチェーン技術に関連する特許情報に対して提案手法を適用し、その結果を用い大企業との差異化を図ることができる新たな中小企業向けビジネス領域の考案を行う。ブロックチェーン技術は実用化が加速している技術領域であり、管理人を必要としない信頼性の高い分散型台帳管理システムとして、決済サービスなどの金融分野をはじめ他の分野での応用が期待されている (Zhao et al., 2016)。そのため新しいビジネス領域の設計概念が生まれる可能性が高い技術領域と考えられ、本研究の題材として適する。

4・2 使用する word2vec モデル

Python の自然言語処理用ライブラリ gensim (Řehůřek and Sojka, 2011) に含まれる word2vec を使用し、コーパスとして英語版 Wikipedia の全記事を採用して学習を行ったモデルを用いる。学習方法には Continuous Bag-of-Word Model (C-BOW) を用い、単語ベクトルの次元数を 200、学習に使う前後の単語数を 15、使用回数が 20 回未満の単語を無視することとする。

4・3 使用した特許情報

ステップ (1) に対応する特許情報の収集には、WIPO (World International Patent Organization) 等の複数のデータベースを用い、一特許分類である CPC および INPADOC (International Patent Documentation) ファミリーからの考察を行った。検索のキーワードとしては様々なものが考えられるが、今回の実施例の対象であるブロックチェーンとスポーツビジネスに直接対応する「blockchain」「sports」の 2 語を設定した。その結果、特許要旨 45 件を収集した。その後、ステップ (2) の手順に従うことで収集した特許要旨は 154 文に分割された。

4・4 抽象概念の階層化と抽象度の設定の結果

ステップ (3) の処理により描画したデンドログラムを図 3 に示す。図 3 内における横軸は文、すなわち抽象概念を表しており、縦軸はコサイン距離を用いた概念間の距離となっている。ステップ (4) における抽象度の設定では、3.5 節で述べた概念のデンドログラムをもとに、クラスタに対応する抽象概念の解釈の難易度と概念の分類数を考慮し、0.87 を閾値 (図 3 内の破線) と決定した。これにより、表 1 に示す 16 個の抽象概念を抽出した。

Table 1 Abstract concepts that collectively represent textual information of corresponding clusters.

Concept 1	Structure of glasses
Concept 2	Electric power (bicycle, motor, solar cell)
Concept 3	Activity units
Concept 4	Providing video contents
Concept 5	Sports game broadcast and result prediction
Concept 6	Broadcasting, authentication of personal information
Concept 7	Motion analysis in sports, health, transaction management, chemistry, etc.
Concept 8	Communication and data storage between devices
Concept 9	Encryption of data communication
Concept 10	IoT devices
Concept 11	Payment and settlement
Concept 12	Sports lottery and lottery tickets
Concept 13	Healthcoin (crypt asset) platform
Concept 14	Gambling and data communication platforms
Concept 15	Raising health awareness
Concept 16	Artificial intelligence technology

Fig. 4 Novelty Potential Matrix and Void Matrix. Case 1 represents the combination of “Concept 2 : Electric power” and “Concept 10 : IoT devices”. Case 2 represents the combination of “Concept4 : Providing video contents” and “Concept16 : Artificial intelligence technology”

4・5 新規性ポテンシャル・Voidマトリクスの作成

ステップ (5), ステップ (6) により作成した NP・Voidマトリクスを図4に示す. この2つのマトリクスによると, 例えば, 「概念2: 電気の動力としての使用 (自転車, モーター, 太陽電池)」と「概念10: IoTデバイス」(以降, ケース1と呼ぶ), および, 「概念4: ビデオコンテンツの提供」と「概念16: 人工知能の技術」(以

Fig. 5 The number of solar cells and IoT patents per year from 2010-2019 (entire world). Search terms are (“solar cells” OR “photovoltaic”) AND (“Internet of Things” OR “IoT”).

Fig. 6 The number of video contents and AI patents per year from 2003-2019 (entire world). Search terms are (“broadcasting” OR “video content” OR “media content”) AND (“artificial intelligence” OR “AI”).

降, ケース 2 と呼ぶ) は, 「新規性ポテンシャルが高い」かつ「今まで存在しなかった」設計概念の組み合わせであることがわかる。

4・6 ビジネス領域のアイデアの考案

各概念の組み合わせを検討し, 新たなビジネス領域の検討を行う。図 4 のマトリックスにおいて新規性ポテンシャルが高く, Void となる概念の組み合わせは多数存在している。その中で, 設計者は, ブロックチェーンを用いた新たなビジネス領域の創出を目的に, 新規性やその実現性, 規模, 収益などの観点から最もビジネス領域として妥当である組み合わせを検討した。その結果, 前節のケース 1, ケース 2 の組み合わせが選ばれ, 各組み合わせに対する新しいビジネス領域の考案を行った。

4・6・1 ケース 1: ソーラーパネルと IoT

ケース 1 に対応する「概念 2: 電気の動力としての使用 (自転車, モータ, 太陽電池)」と「概念 10: IoT デバイス」の組み合わせから, スポーツ用の IoT デバイスの動力源として太陽電池を用いるなどのビジネス領域のアイデアが創出された。具体的な新規事業案として, 将来的には, 数十億台の IoT デバイスが屋内太陽電池により動力を得て, 環境情報から H2M (Human-to-Machine) 通信, M2M (Machine-to-Machine) 通信までを提供し, エネルギー消費量や電池の無駄を削減することができると考えられる。したがって, これらの概念の組み合わせは, 技術戦略の柔軟性に富む中小企業が製品やサービスの面で, 他の大企業等の競合他社との差異化を達成する上で効果的であると考えられる。

4・6・2 ケース 2: ビデオコンテンツと AI

ケース 2 に対応する「概念 4: ビデオコンテンツの提供」と「概念 16: 人工知能の技術」の組み合わせから, ブロックチェーン, AI, IoT デバイスを活用した動画プラットフォームを格闘技向けに構築するビジネス領域が創出された。具体的には, 格闘技のトレーナーと選手を映像によって繋ぎ, トレーニング実績やトレーニング回数, 運動能力のレベルに応じた動画や研修プランなどパーソナライズされたコンテンツを提供する事業が考えられる。ブロックチェーンベースのプラットフォームを使用することで個人情報を安全に保存できることや, COVID-19 の現代においてはビデオを用いることでいつでもどこでもトレーナーと繋がることのできるなどの利点を有する。

5. 考察

4章に示した本手法の実施例では、スポーツビジネス研究との融合により、本提案手法に従って作成された2つのマトリクスの結果から着想を得て、新たなビジネス領域の設計概念を生成することができた。2.1節で述べたように、近年のデータサイエンスの発展に伴い、大量の特許情報からアイデア生成を支援する研究が行われている。しかし、従来研究では、特許の文章内容を活用した支援に課題が残されていた(Luo et al., 2018)。本研究では特許の要旨を集め、単語分散表現モデルを用いることにより、非構造データである特許の文章内容を考慮した分析を可能とした。具体的には、設計概念生成の理論であるECFサイクルとNP計量化手法に基づいて、特許に含まれる抽象概念を抽出し、それらの間の距離を測定し、NP・Voidマトリクスを生成してユーザーに提示することで、効率的にビジネス領域のアイデアの着想を行うことができた。この結果は、本手法が特許文章の内容を活用し、新たな設計概念を生成する手法として有効であるということを示している。

4章で用いた特許はスポーツビジネスに限定したものであった。得られた技術領域のアイデアの有効性をより客観的に検証するために、スポーツビジネス以外の特許を含めて、それぞれの技術トレンドを調査した。まず、4・6・1項のケース1に対応する「("solar cells" OR "photovoltaic") AND ("Internet of Things" OR "IoT")」の条件で語句検索を行い、2010～2019年における年度ごとの特許数の推移を表したものを図5に示す。図5より、当該の語句の組み合わせを含む特許は2015年から2018年にかけて上昇していることが分かる。また4・6・2項のケース2に対応する「("broadcasting" OR "video content" OR "media content") AND ("artificial intelligence" OR "AI")」の条件で語句検索を行い、2003～2019年における年度ごとの特許数の推移を表したものを図6に示す。こちらも、当該語句の組み合わせを含む特許は特に2017年から2019年にかけて上昇していることが分かる。これらの結果から、ケース1およびケース2に関係する特許出願数は近年増加傾向にあることが判明した。このことより考案したビジネス領域は、スポーツビジネス以外でも発展途上の新規性の高い技術分野を利用したものであると考えることができ、提案手法は新規性ポテンシャルの高い設計概念の生成に有用であることを示している。

本研究の提案手法では、ステップ(4)「抽象度の設定」を人間がクラスタの意味を解釈して判断している。件数が増えた場合には何らかの形で計算機支援が必要と考えられる。例えば、グループ内の重心に近い文を代表文として取り上げるなどの方法が考えられる。その検討は今後の課題である。また、本研究で提案する手法での支援の対象は、新規性ポテンシャルが高いとされるVoidを提示することまでであり、図2のステップ(7)に対応する新しい実体概念を生成する操作は直接の支援の対象となっていない。図4に示した抽象概念の組み合わせの中には、「概念1：メガネの構造機構」と「概念12：スポーツ賭博・宝くじ」の組み合わせのように、概念間の距離が遠すぎて、対応する実体概念を着想することが難しい組み合わせも多数存在している。そのような場合でも、設計概念生成を効率的に行うことが求められる。これに向けては、抽象概念と実体概念の関係を踏まえて、対応するVoidに近い既存の実体概念を抽出・要素化し、適宜組み合わせで新たな実体概念を生成する、などの一連の概念操作の支援が必要である。本研究で提案する手法では、特許情報に基づいて抽象概念と実体概念の対応関係を生成した上で、単語分散表現モデルによる距離の導入を行っており、それに向けた基盤ができていると考えられるものの、具体的な概念操作の形式化と支援方法の確立は今後の課題である。また、概念空間を十分に構造化することにより、例えば、先行研究(野間口他, 2017)で提案しているようなTRIZ(発明的問題解決理論)に代表される既存の発想支援手法をより有効に用いることができると考えられ、その検討も今後行う必要がある。

本研究では、英語で記述された特許文書を扱ったが、日本語の特許文書においても要約が存在し、また、日本語のwikipediaで学習したword2vecモデル(鈴木他., 2016)を利用することで、同様のアプローチを用いることができると考えられる。

6. 結言及び今後の展望

本研究では技術的知識が集約された特許情報に注目し、著者らのグループで提案されているECFサイクルの設計概念生成手法に基づいて、設計概念を生成する手法を提案した。本手法の実施例ではブロックチェーンを用いたスポーツビジネス領域の特許を題材として取り上げ、2つの具体例を抽出し本手法の有効性を確認した。今後は自然言語処理の特徴を活かし、さらに大規模な特許データ解析を行うことにより、設計概念の探索範囲を大き

くすることや抽象概念の選択段階における指標について調査を行うことで本手法の実用性を高めることを目指す予定である。

文 献

- Abdelgawad, L., Kluegl, P., Genc, E., Falkner, S. and Hutter, F., Optimizing neural networks for patent classification, Proceedings of Joint European Conference on Machine Learning and Knowledge Discovery in Databases (2019), pp. 688-703.
- Albert, M. B., Avery, D., Narin, F. and McAllister, P., Direct validation of citation counts as indicators of industrially important patents, Research policy, Vol. 20. No. 3 (1991), pp. 251-259.
- Bird, S., NLTK: the natural language toolkit, Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions (2006), pp. 69-72.
- Kuznets, S., Inventive activity: Problems of definition and measurement, The Rate and Direction of Inventive Activity: Economic and Social Factors, Princeton University Press (1962), pp. 19-52.
- Lee, C., Jeon, D., Ahn, J. M. and Kwon, O., Navigating a product landscape for technology opportunity analysis: A word2vec approach using an integrated patent-product database, Technovation, Vol. 96 (2020), pp. 102140.
- Lee, S., Yoon, B. and Park, Y., An approach to discovering new technology opportunities: Keyword-based patent map approach, Technovation, Vol. 29, No. 6-7 (2009), pp. 481-497.
- Lee, Y., Kim, S. Y., Song, I., Park, Y. and Shin, J., Technology opportunity identification customized to the technological capability of SMEs through two-stage patent analysis, Scientometrics, Vol. 100. No. 1 (2014), pp. 227-244.
- Leydesdorff, L., Kushnir, D. and Rafols, I., Interactive overlay maps for US patent (USPTO) data based on International Patent Classification (IPC), Scientometrics, Vol. 98. No.3 (2014), pp. 1583-1599.
- Luo, J., Song, B., Blessing, L. and Wood, K., Design opportunity conception using the total technology space map, Artificial Intelligence for Engineering Design, Analysis and Manufacturing (AIEDAM), Vol. 32. No. 4 (2018), pp. 449-461.
- Mikolov, T., Chen, K., Corrado, G. and Dean J., Efficient estimation of word representations in vector space, arXiv preprint arXiv:1301.3781 (2013).
- 野間口大, 河原貴大, 正田浩暉, 藤田喜久雄, 設計概念生成のための新規性ポテンシャルの考え方と単語類似度計測によるその評価法の提案, 日本機械学会論文集, Vol. 86, No. 890 (2020), DOI:10.1299/transjsme.20-00006.
- 野間口大, 中切拓, 藤田喜久雄, 外在化・分類操作・Void 発見のサイクルと発明的問題解決理論の連係による発想支援手法, 日本機械学会論文集, Vol. 83, No. 853 (2017), DOI:10.1299/transjsme.17-00122.
- Řehůřek, R. and Sojka, P., Gensim—statistical semantics in python, Retrieved from gensim.org (2011).
- Rousseeuw, P. J., Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, Journal of Computational and Applied Mathematics, Vol 20 (1987), pp.53-65.
- Schakel, A. M. J. and Wilson, B. J., Measuring word significance using distributed representations of words, arXiv preprint arXiv:1508.02297 (2015).
- Shibata, M., Ohtsuka, Y., Takahashi, M. and Okamoto, K., Advanced FPGA technology trend based on patent analysis with link mining, Proceedings of 2018 International Conference on Electronics Packaging and iMAPS All Asia Conference (ICEP-IAAC 2018) (2018), pp.147-151 (Session Invited).
- 鈴木正敏, 松田耕史, 関根聡, 岡崎直観, 乾健太郎, Wikipedia 記事に対する拡張固有表現ラベルの多重付与, 言語処理学会第 22 回年次大会発表論文集, (2016), A5-2.
- Taura, T. and Nagai, Y., A systematized theory of creative concept generation in design: first-order and high-order concept generation, Research in Engineering Design, Vol. 24, No. 2 (2013), pp. 185-199.
- Yang, C. H., Constructing design patent map and design strategy based on two perspective of classifying methodology, Proceedings of the 5th International Congress of International Association of Societies of Design Research (IASDR) (2013), 2059-1.
- 吉川弘之, 一般デザイン学 (2020), 岩波書店.
- Zhao, J. L., Fan, S. and Yan, J., Overview of business innovations and research opportunities in blockchain and introduction to the special issue, Financial Innovation, Vol. 2 (2016), 28.

References

- Abdelgawad, L., Kluegl, P., Genc, E., Falkner, S. and Hutter, F., Optimizing neural networks for patent classification, Proceedings of Joint European Conference on Machine Learning and Knowledge Discovery in Databases (2019), pp. 688-703.
- Albert, M. B., Avery, D., Narin, F. and McAllister, P., Direct validation of citation counts as indicators of industrially important patents, Research policy, Vol. 20. No. 3 (1991), pp. 251-259.
- Bird, S., NLTK: the natural language toolkit, Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions (2006), pp. 69-72.
- Kuznets, S., Inventive activity: Problems of definition and measurement, The Rate and Direction of Inventive Activity: Economic and Social Factors, Princeton University Press (1962), pp. 19-52.
- Lee, C., Jeon, D., Ahn, J. M. and Kwon, O., Navigating a product landscape for technology opportunity analysis: A word2vec approach using an integrated patent-product database, Technovation, Vol. 96 (2020), pp. 102140.
- Lee, S., Yoon, B. and Park, Y., An approach to discovering new technology opportunities: Keyword-based patent map approach, Technovation, Vol. 29, No. 6-7 (2009), pp. 481-497.
- Lee, Y., Kim, S. Y., Song, I., Park, Y. and Shin, J., Technology opportunity identification customized to the technological capability of SMEs through two-stage patent analysis, Scientometrics, Vol. 100. No. 1 (2014), pp. 227-244.
- Leydesdorff, L., Kushnir, D. and Rafols, I., Interactive overlay maps for US patent (USPTO) data based on International Patent Classification (IPC), Scientometrics, Vol. 98. No.3 (2014), pp. 1583-1599.
- Luo, J., Song, B., Blessing, L. and Wood, K., Design opportunity conception using the total technology space map, Artificial Intelligence for Engineering Design, Analysis and Manufacturing (AIEDAM), Vol. 32. No. 4 (2018), pp. 449-461.
- Mikolov, T., Chen, K., Corrado, G. and Dean J., Efficient estimation of word representations in vector space, arXiv preprint arXiv:1301.3781 (2013).
- Nomaguchi, Y., Kawahara, T., Shoda, K. and Fujita, K., Proposal of the novelty potential for design concept generation and its assessment method with word similarity, Transactions of the JSME (in Japanese), Vol. 86, No. 890 (2020), DOI:10.1299/transjsme.20-00006.
- Nomaguchi, Y., Nakigiri, T. and Fujita, K., Concept generation support with linkage of externalization-classification-finding void cycle and theory of solving inventive problems, Transactions of the JSME (in Japanese), Vol. 83, No. 853 (2017), DOI:10.1299/transjsme.17-00122.
- Řehůřek, R. and Sojka, P., Gensim—statistical semantics in python, Retrieved from gensim.org (2011).
- Rousseeuw, P. J., Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, Journal of Computational and Applied Mathematics, Vol 20 (1987), pp.53-65.
- Schakel, A. M. J. and Wilson, B. J., Measuring word significance using distributed representations of words, arXiv preprint arXiv:1508.02297 (2015).
- Shibata, M., Ohtsuka, Y., Takahashi, M. and Okamoto, K., Advanced FPGA technology trend based on patent analysis with link mining, Proceedings of 2018 International Conference on Electronics Packaging and iMAPS All Asia Conference (ICEP-IAAC 2018) (2018), pp.147-151 (Session Invited).
- Suzuki, M., Matsuda, K., Sekine, S., Okazaki, N. and Inui, K., Multiple assignment of extended named entity labels for Wikipedia articles, Proceedings of the Twenty-second Annual Meeting of the Association for Natural Language Processing, (2016), A5-2 (in Japanese).
- Taura, T. and Nagai, Y., A systematized theory of creative concept generation in design: first-order and high-order concept generation, Research in Engineering Design, Vol. 24, No. 2 (2013), pp. 185-199.
- Yang, C. H., Constructing design patent map and design strategy based on two perspective of classifying methodology, Proceedings of the 5th International Congress of International Association of Societies of Design Research (IASDR) (2013), 2059-1.
- Yoshikawa, H., General Design Theory (2020), Iwanami Shoten, (in Japanese).

Zhao, J. L., Fan, S. and Yan, J., Overview of business innovations and research opportunities in blockchain and introduction to the special issue, Financial Innovation, Vol. 2 (2016), 28.