

Title	Ai計算による「組織対組織の产学連携」への試み
Author(s)	開本, 亮; 難波, 英嗣; 杉山, 典正
Citation	年次学術大会講演要旨集, 40: 797-800
Issue Date	2025-11-08
Type	Conference Paper
Text version	publisher
URL	https://hdl.handle.net/10119/20177
Rights	本著作物は研究・イノベーション学会の許可のもとに掲載するものです。This material is posted here with permission of the Japan Society for Research Policy and Innovation Management.
Description	一般講演要旨



Ai 計算による「組織対組織の产学連携」への試み

○開本 亮 (大阪工業大学), 難波英嗣 (中央大学), 杉山典正 (大阪工業大学)

Akira.hirakimoto@oit.ac.jp

1. はじめに

発表者らは、昨年の本学会にて、Ai 計算による大学研究者のボーア・エジソン・パスツール分類を行い、ボーア成分、エジソン成分、パスツール成分からなる 3 次元 BEP ベクトルを大学研究者毎に算出し、そのベクトル演算によって、共同成果の予測計算が一定程度の精度で可能なことを示した。なお、3 次元の座標は、論文分類座標、特許分類座標、科研費分類座標であり、それぞれボーア成分、エジソン成分、パスツール成分に対応する。そして、これを発展させて、関西圏の 23 大学が参画する KSAC (<https://ksac.site/>) において、大学研究者を対象とした共同成果の予測計算のシステムを稼働させる準備を行なせている。

しかしながら、上記の Ai 計算は学術論文に限定されているため、大学と企業との産学連携には適用が困難であった。

そこで、Ai 計算を企業発特許も対象とする開発も進行させており、今般、最初の計算結果として、京都大学と島津製作所の結果が得られたので、以下に報告する。

2. 学術論文を対象とした共同成果の予測計算

昨年の本学会にて報告した共同成果の予測計算を図 1 に示す（詳細は昨年度予稿集 1G23 等を参照 [1][2]）。即ち、研究者 Y と研究者 O を対象として、Ai 計算によって両者の 3 次元 BEP ベクトルを算出し、その成す角度 θ について $\sin^2 \theta$ （ヘテロファクター）を計算する。これは両者の将来における論文創出の可能性を示す 3 次元スカラー量である。次に、両者の論文の相乗平均を計算する。これは両者の過去における論文創出の共通基盤を示す 3 次元スカラー量である。最後に両スカラー量を乗じることで、3 次元上、即ち、論文分類座標、特許分類座標、科研費分類座標上に、共同成果の創出確率が示される。

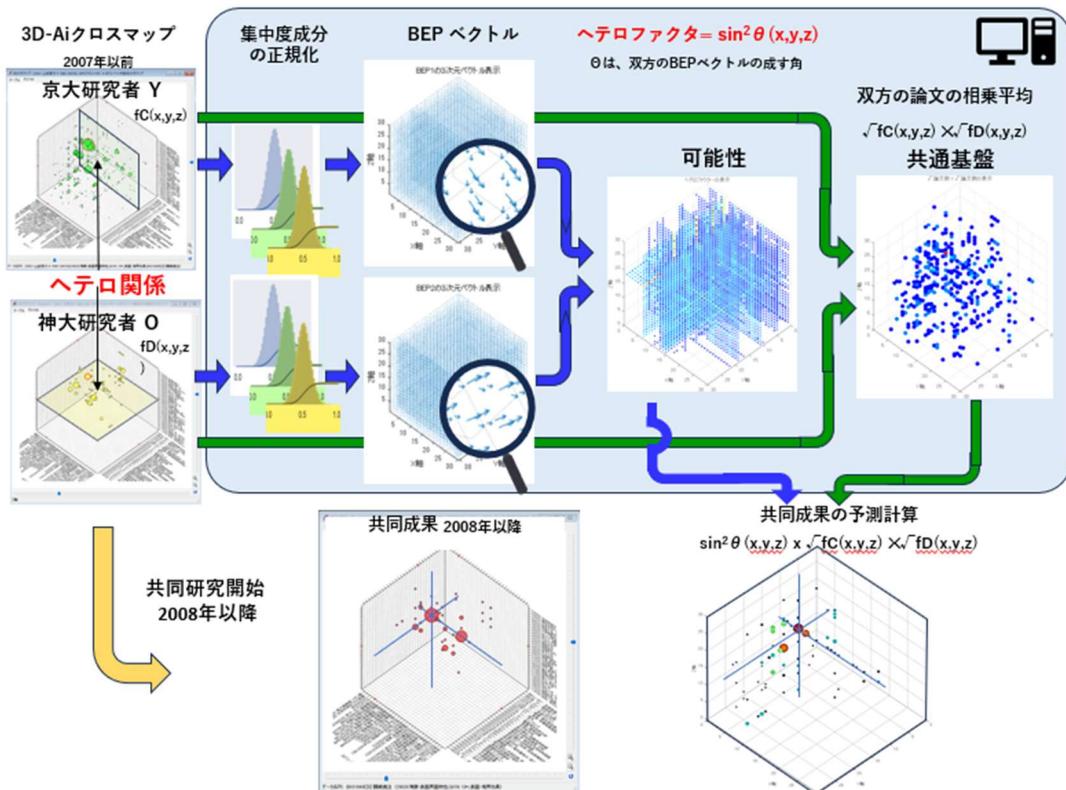


図1 共同成果の予測計算

3. 企業発特許への拡張の問題点とその解決手段

まず、拡張しない大学発論文のみの場合では、3次元 BEP ベクトルに算出するために、論文のアブストラクトを入力して Deep Learning によって特許分類を推定計算する「①特許分類器」と、同様に科研費分類を推定計算する「②科研費分類器」が必要であり、これらは共同発表者である難波が開発したシステムである。なお論文分類は、科学技術振興機構（JST）による論文分類をそのまま利用する。

次に、企業発特許に拡張する場合は、特許のアブストラクトを入力して論文分類を推定計算する「③論文分類器」と、同様に科研費分類を推定計算する「④科研費分類器」が必要となる。ここで、科研費分類は他の分類に比較して分類数が1桁以下の300程度であり、特許のアブストラクトを入力しても、計算精度に問題ないレベルであるので、「②科研費分類器」＝「④科研費分類器」とすることができ、必須のものは「③論文分類器」となる。なお特許分類は特許庁（JPO）による特許分類をそのまま利用する。

さてここで「③論文分類器」に関して、企業発特許は30条適用出願（新規性喪失の例外適用出願）の割合が約0.5%と非常に少ないため、特許に論文を直接的に紐付ける教師データが不足しており、加えて特許独特の文言表現等により Deep Learning システムの開発は難度が高かった。

そこで、発表者は教師データがなくとも特許から論文分類を推定計算可能な「③論文分類器」を図2および3に示す構成で考案した[3]。

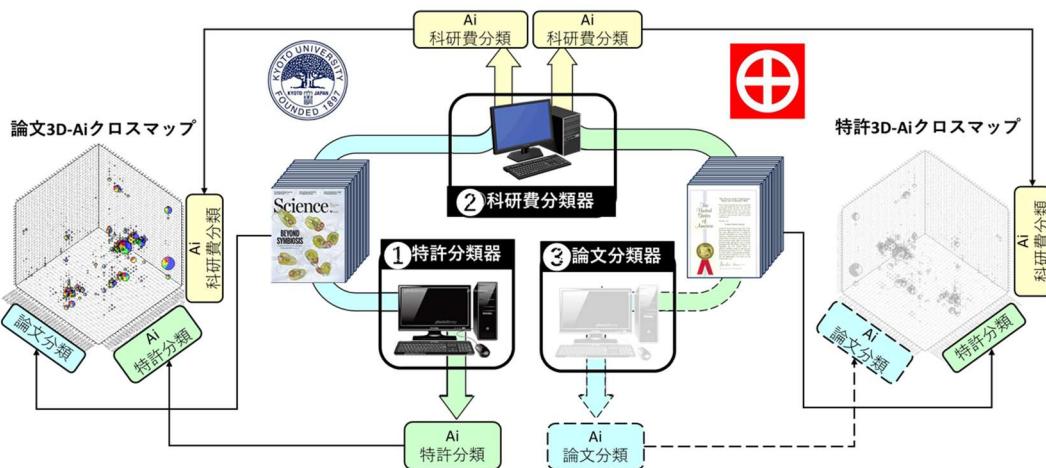


図2 2次元分類共通座標の取得

まず図2に示すように、産学連携を検討する対象の大学発論文群（例えば京都大学）と、企業発特許群（例えば島津製作所）を予め決めた上で、企業発特許群に対して、特許庁が付与した特許分類をそのまま利用し、更に特許のアブストラクトを入力して「②科研費分類器」により科研費分類を算出する。また従来と同様に、大学発論文に対して、「①特許分類器」により同分類を算出し、更に「②科研費分類器」により科研費分類を算出する。そうすると、大学発論文群と企業発特許群について、図3の左右に示す縦軸が特許分類、横軸が科研費分類の2次元分類共通座標を取得することができる。

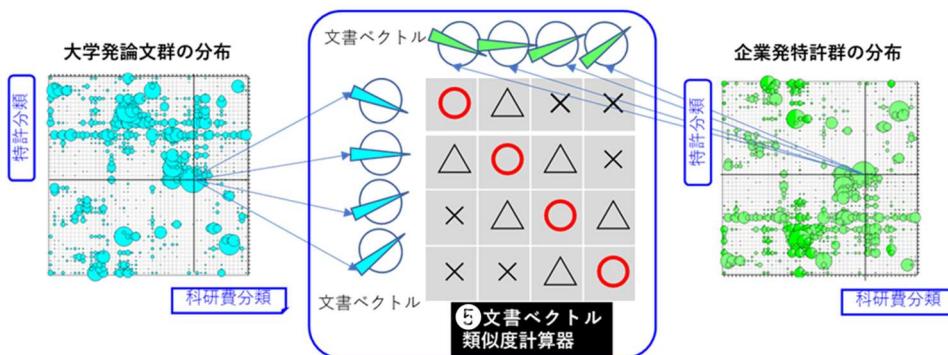


図3 類似度計算による論文に対応する特許の特定

そして、上記2次元分類共通座標の分布において、同一座標の大学発論文群と企業発特許群は、一定程度で類似するものと考えられる。2次元の同一座標は同一分類となるからである。

そこで、「⑤文書ベクトル・類似度計算器」により、同一座標の大学発論文群と企業発特許群の類似度を文書ベクトル等で計算すると、特定の企業発特許に最も類似する、特定の大学発論文（○）を決定でき、その大学発論文にJSTが付与した論文分類が判明する。これを特許に対する論文分類とすればよいのである。

したがって、図2および図3を一体として捉えると、特許に論文分類を付与するシステムになっているので、企業発特許に教師データがなくとも、「③論文分類器」として動作する。勿論、類似度が低ければ類似しないとして棄却を行い、極めて高ければそのまま产学研連携の候補となり得るのである。

このシステムの最初のトライアルとして、図4に示すように、京都大学論文13,178件（2018年～2022年）と島津製作所特許1,369件（2019年～2023年）の類似度計算を行った。ここで、2次元分類共通座標では、一つの論文や特許に対して分類数が複数の場合、複数座標に現れるので、計算の入力対象の論文テキストデータは28,320件、特許テキストデータは5,103件となっている。

仮に総当たりで類似度計算を行うと、 $28,320 \times 5,103 = 1$ 億4,000万回の計算が必要となるが、本システムでは同一座標のペアを計算すればよいので50万回の計算で完了する。したがって、総当たり計算では37時間30分かかっていたが、本システムでは8分という短時間で完了し、漸く実用レベルで利用できるものとなつた。

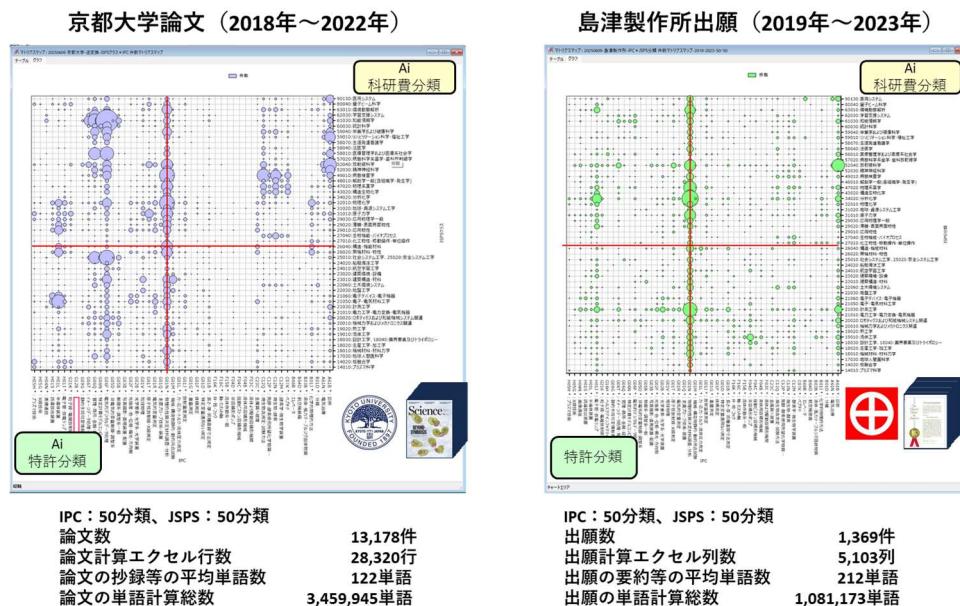


図4 組織対組織の产学研連携への応用 京都大学と島津製作所

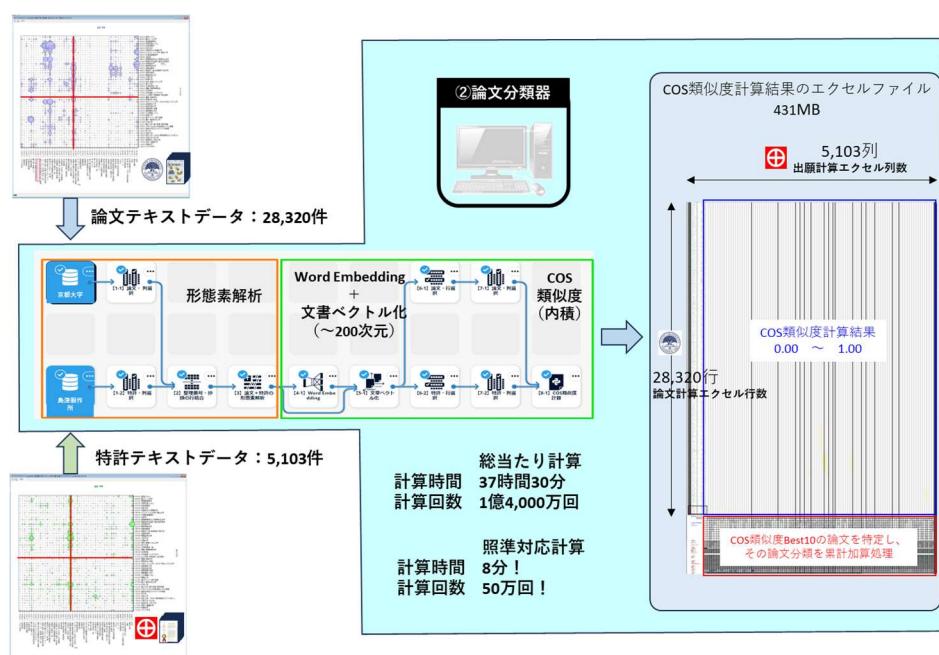


図5 組織対組織の产学研連携への応用 具体的計算方法

4. 計算結果の第一報

本システムによって計算した京都大学と島津製作所の類似度計算の第一報（一部）を表1に示す。左から第1列が島津製作所の特許であり、第2列がそれに最も類似する京都大学の論文であり、第3列が科学技術推進機構が付与した論文分類と、本システムによる類似度である。「？」等を付した論文分類は計算精度に疑問が残るものである。表1の第2列を見ると、共通座標における類似論文候補数が10件以上では精度が高く、「？？」を全て付したものは、同候補数が2件のみの場合であった。

なお、表1の中で、30条適用の島津製作所の特許出願は第4行の「質量分析データ処理方法、…」であり、その論文に対して科学技術振興機構が付与した論文分類は、EG02020G：微生物検査法、EA03020V：生物物理的研究法、JE08000Z：人工知能であった。本システムによる計算結果は、JE08000Z：人工知能（0.201）、JE04020T：医用画像処理（0.198）、FF05020Z：木材の性質・構造（0.105）であり、JE08000Z：人工知能が一致した。

表1 組織対組織の产学連携への応用 結果の第一報（一部）

島津製作所 出願番号 30条適用出願	COS類似度計算結果 類似度最大の著者・論文	COS類似度計算結果 類似度Best10の論文分類とその累計加算類似度
島津製作所 特願2019-067448 「食品の品質判定方法、及び、食品品質判定装置」	京都大学 農業研究科 ○○○○ 「人工知能を用いて食品内部のミクロ構造から品質を予測する」 共通座標における類似論文候補数：10件以上	COS類似度合計が0.5以下のため、参考情報 FJ01040F 食品の品質 0.206 FJ01010Y 食品一般 0.104 AE01000T 心理学一般 0.104 ??
島津製作所 特願2019-089135 「短鎖塩素化バラフィンの同族体の検出方法」	京都大学 防災研究所 ○○○○ 「ドローン熱画像撮影による砂州における伏流水湧出場の検出」 共通座標における類似論文候補数：2件のみ	RC12020M 河川調査・計画 0.554 ??? FF03040H 森林保育 0.446 ???
島津製作所 特願2019-147020 「X線透視撮影装置」	京都大学 附属病院 ○○○○ 「照射内コーンビームコンピュータ断層撮影画像に関する実際の送達量計算：ファントム研究」 共通座標における類似論文候補数：10件以上	JE04020T 医用画像処理 0.292 GA03000B 法医学 0.109 ? GA05020L 医療用機器装置 0.104
島津製作所 特願2019-157607 「質量分析データ処理方法、質量分析データ処理システム、及びプログラム」	京都大学 附属病院 ○○○○ 「病理画像癌種別領域分割のための癌種比率を活用した学習手法」 共通座標における類似論文候補数：10件以上	JE08000Z 人工知能 0.201 JE04020T 医用画像処理 0.198 FF05020Z 木材の性質・構造 0.105 ???
島津製作所 特願2020-155761 「うつ秒の診断補助方法」	京都大学 附属病院 ○○○○ 「うつ病の分子基盤：エビジェネティック制御における性関連差異の意味」 共通座標における類似論文候補数：10件以上	GR03000A 精神障害 0.299 GR02000T 精神障害の診断 0.103 GN02000H 神経系の診断 0.102
島津製作所 特願2020-175127 「質量分析を用いた試料分析方法及び試料分析システム」	京都大学 理学研究科 ○○○○ 「LA-ICP-MSによる鉛フリーはんだチップ中の主成分および微量元素の定量分析」 共通座標における類似論文候補数：10件以上	COS類似度合計が0.5以下のため、参考情報 CC03040V 金属、合金の物理分析 0.102 ? CC04014L 液体クロマトグラフィ 0.101 CC01040H 抽出 0.101
島津製作所 特願2021-127548 「アミノ酸の分析方法」	京都大学 医学研究科 ○○○○ 「MALDI-TOF/MSを用いた尿中リン脂質とリソリル脂質の抽出と分析の比較評価」 共通座標における類似論文候補数：10件以上	COS類似度合計が0.5以下のため、参考情報 CF10080G ベプチド 0.199 EF04010S 生体膜一般 0.105 ? CC08080W 水中の汚染物質の各種 0.103 ?

5. 今後の展開

第一報の結果から、精度向上のため、共通座標における類似論文数を増加させることが第一義的に重要であることがわかる。そこで、KSAC 参画 23 大学の論文約 29 万件を対象とするように、約 20 倍のスケールアップを目指す。また本予稿では言及しなかったが、同義語・類義語辞書の整備、文章ベクトル化における最適次元数の探索、特許テキストの単語数拡張等も行う予定である。

最後に、本システムが完成すれば、従来、互いの情報不足で円滑な運営が十分でなかった产学連携について、共同成果の予測計算が可能となり、組織対組織の定量的・俯瞰的・戦略的な展開が可能となり、产学連携によるイノベーション促進に大きな貢献ができると期待される。

謝辞

本発表のデータ分析には、株式会社 NTT データ数理システム（特にデータ分析プラットフォーム「Alkano」の適用について北沢芳明氏）、株式会社ジー・サーチ、インパテック株式会社のご協力を得ました。ここに深く感謝いたします。また本研究は科学研究費（24K05100）を受けて行われたものです。

参考文献

- [1] 開本 亮, Ai クロスマップによる戦略的产学連携の試み 研究技術計画, 35, 329-338 (2020)
- [2] 開本 亮, 特開 2025-069930 「情報処理装置、情報処理方法、およびプログラム」
- [3] 開本 亮, 特開 2024-123312 「プログラム、逆クロス分類装置、およびテキストの分類方法」