

機械学習を用いた学際分野の俯瞰的分析：学術論文の文章と図の分類と可視化

菅原 裕輝（人文学研究科）

森田 邦久（人間科学研究科）

大塚 淳（ZEN 大学 知能情報社会学部）

中島 悠太（D3 センター）

1 研究の背景

現代の学術研究は専門分化（細分化）と学際化が同時に進行し、知識は深く掘り下げられる一方で分野間の相互交流によって新たな領域が生み出されている。伝統的に学問の体系は、一元的・階層的なモデルで理解されてきた（例：コントの学問序列、論理実証主義による全科学の統一）。これらの統一・還元的な見方は知の構造を簡明に示す利点があるものの、多元的で相互循環的に発展する現代の知のあり方を十分に説明しきれない可能性がある。明治期の西周『百学連環』においては、学問同士を有機的に関連させる環状モデルが示されたが、21 世紀の現在では学問領域は一層多様化・変容し、その構造はかつてない複雑さを帯びている。このような状況下で、学術分野全体の俯瞰的理解（ひいては、知の構造の解明）を試みるには、新たな分析視座と手法が求められる。

近年のデータサイエンスや機械学習の発展により、大規模なテキストや画像データを計算機で分析し、そこから知見を得ることが可能になってきた。学術論文のデータも例外ではなく、デジタル人文学の文脈では、膨大な論文テキストや引用関係を解析して科学知識の構造を明らかにする試みが現れている。本研究もこの潮流に位置づけられ、機械学習技術を用いて学際的な研究領域を俯瞰し、専門分野の細分化と学際的統合のダイナミズムを明らかにすることを目指す。

2 研究の目的

本研究の目的は、機械学習を活用して学術論文の文章テキストと図データを分類・可視化し、学際領域における知の構造を包括的に解明することである。具体的には、異なる学問分野間で知識がどのように関連づき、融合し、新たな研究領域や概念が生み出されているのかをデータ駆動型で検出・可視化する点にある。特に、学術論文に含まれるテキスト情報（タイトル・要旨・本文など）と図情報を統合的に分析し、従来は個別に扱われてきた知識の言語表現と視覚表現の両面から、学際分野に特有の特徴や傾向を明らかにすることを目指した。これにより、従来型の階層モデルでは捉えがたい学問分野間の横断的な知識の流動や構造を明らかにし、学際研究を理解する新たな視座を提供する。

3 研究の方法

本研究では、知識構造の可視化を試みた。とりわけ、進化生物学分野や神経科学分野、科学哲学分野、科学技術社会論分野といった学際性の高い学術領域の論文を分析対象とし、データの収集を進めてきている。また、画像データの分析のケーススタディとして、論文の図を扱うための前段階として、意図や感情が読み取りやすいソーシャルメディア上の画像データを分析

対象とした。

3.1 テキストデータの分析手法

テキストデータに対しては、自然言語処理とトピックモデルによる分析を行った。まず語彙頻度の集計やワードクラウドの生成によってデータ全体の傾向を把握した上で、BERT に代表される事前学習言語モデルを用いて論文テキストをベクトル表現に変換した。高次元ベクトルで表された文書間の類似度関係は、次元削減アルゴリズム UMAP によって可視化し、文書群の分布やクラスター構造を直観的に把握できるようにした。さらに、これらのベクトル表現に基づき HDBSCAN によるクラスターリングを適用し、トピックモデル手法 BERTopic を用いて文脈的に一貫したトピックの抽出を行った。BERTopic は文章埋め込みとクラスターリングにより代表的な話題を自動抽出する手法であり、各クラスターに含まれる論文の主要キーワード群から意味的に特徴づけられたトピックを生成する。本研究では、これにより約 40 のトピックを同定し、それぞれのトピックに代表語（キーワード）の集合を割り当てた。また、抽出されたトピックが時系列的にどのように変化してきたかを分析するため、各年に発表された論文中のトピック出現割合を算出し、トレンド分析を行った（論文投稿の準備中）。

3.2 画像データの分析手法

学術論文中の図表あるいは関連する画像データに対しては、画像内容をテキストで記述する生成 AI モデルを用いた分析を試みた。本研究では論文の図を分析する前段階として、意図や感情の解釈がしやすいソーシャルメディアの画像の分析を行った（雑誌論文 1）。具体的には、視覚と言語の両領域で事前学習されたマルチモーダルモデルである BLIP（Bootstrapping Language-Image Pre-training）を利用し、画像から自動的にキャプション（内容説明文）を生成した。これは画像に潜む特徴を人間の言葉で表現させることで、図像データをテキストデータと同様に扱い解析することを可能にするアプローチであり、ある種の説明可能 AI（Explainable AI）の手法と言える。生成された画像キャプションについてテキストマイニングを行い、頻出する語句や内容の傾向を分析することで、図像情報に内在するテーマや情動的含意を読み解いた。例えば、異なるデータソース（学術誌やプラットフォーム）間で生成キャプションを比較することで、画像が一般的にどのように認識され、どのような情動反応を引き起こしうるかの違いを可視化した。

なお、本研究では今後の発展に向けて、引用関係データを組み込んだグラフニューラルネットワーク（GNN）による論文間関係性の解析や、畳み込みニューラルネットワーク（CNN）による学術図表の自動分類も検討している。現段階では未導入で

あるものの、こうしたネットワーク分析や高精度な画像分類技術を組み合わせることで、テキスト・図・ネットワーク情報を統合したより包括的な知識構造の俯瞰分析へと発展させる予定である。

4 研究成果

本研究の手法を用いて、大量の学術情報から学際的な知識構造を可視化することに成功した。例えば、哲学分野のトップジャーナルであり学際的な研究を多く掲載する学術誌 *Synthese* の論文タイトルデータ（1997～2025年の論文タイトル計7284件）にトピックモデル（BERTopic）を適用した結果、抽出されたトピックの内容から哲学内部の伝統的テーマ（例：論理学、認識論）に加えて、数理学、認知科学、神経科学といった科学分野にまたがる学際的なテーマが浮かび上がった。図1は抽出されたトピック間の関係を可視化したものである。各点は論文トピックを表し、その距離はトピック間の語彙意味の近さを示している。図中、左上に位置するトピック群は哲学固有のテーマ（例：形而上学、倫理学などの伝統的トピック）であり、右下に位置するトピック群は科学的テーマ（例：統計学、物理学、神経科学など）に対応していた。両者の中間には融合的性格を持つトピックが分布し、哲学と自然科学の境界が連続的であることが視覚的に示された。言い換えれば、哲学分野の研究が科学的知見を積極的に取り入れて発展しており、学問領域間の明確な断絶ではなくグラデーション的の連関関係として把握できることが明らかとなった。

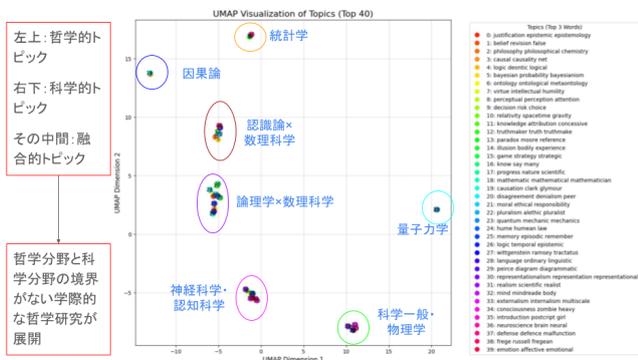


図1: トピックのUMAP視覚化

次に、各トピックの年代別動向を分析した結果、学際トピックの出現頻度には時期による変遷が見出された。図2は1997年から2020年代にかけての論文集合における主要トピックの時間的推移を示している。一例として、神経科学に関連するトピック（図2中のTopic 36）は2000年代半ばから2010年代前半にかけて相対的に高い存在感を示し、その後減少した。一方、認知科学に関連するトピック（Topic 8）は2010年代後半以降に台頭し、哲学誌上での注目度が高まっている。また、数理学に関するトピック（Topic 5）は2010年前後に顕在化している。これらの結果から、対象とした哲学分野では時代とともに科学との接点が変わってきたこと、すなわちある時期には神経科学との交流が盛んであったのが、その後は認知科学との結びつきが強まるなど、学際的関心の推移が定量的に捉えられた。

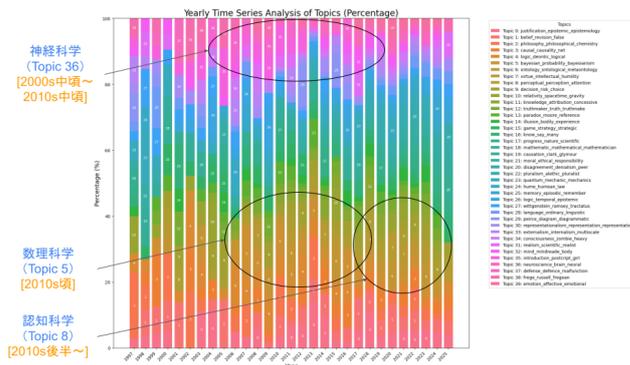


図2: トピックの時系列分析

以上の分析より、専門分化が進む中でも学際的な知の循環が生じていることが示唆された。哲学分野の事例では、従来の還元的・階層的モデルでは捉えきれない多層的かつ循環的な知識のダイナミクスが確認できる。実際、*Synthese* 誌の論文は哲学固有のテーマを保持しつつ、数学・認知・神経科学といった周辺領域へと展開していることが示された。このことは、細分化による知の断片化が学際的連関によって補完され、知識同士が循環的に関連づく可能性を示す。換言すれば、西周の提唱した「百学連環」の思想を現代的にアップデートし、学問領域の枠組みを超えた知の構造を描き出せることを示したと言える。

さらに、本研究の枠組みは他の学際的課題にも適用可能であることを示すため、デジタル人類学の研究事例に当てはめた分析も実施した（雑誌論文1）。具体的には、津波被災地域への移住に関するソーシャルメディア上の発信（テキストおよび写真データ）を対象とし、複数プラットフォーム間での情報内容の違いを抽出した。テキスト投稿のトピック分析では、YouTube・Facebook・Instagramの各プラットフォームで移住に関する主要話題が異なり、移住者が媒介ごとに異なる観点から情報発信していることが明確となった。例えばFacebookでは地域社会の共同活動や復興プロジェクトに関する話題が目立つのに対し、Instagramでは自然環境下での生活や娯楽的要素に焦点を当てた発信が多い、といった傾向が見られた。さらに画像データに対してBLIPモデルで生成したキャプションの分析から、プラットフォーム間で強調される情動的側面の違いが浮かび上がった。Facebook投稿の画像では「人々が協力して作業している様子」を記述するキャプションが多く、連帯感・協働といった社会的情動を喚起する内容が多く観察された。一方、Instagram投稿の画像では「浜辺で跳ねる子供」「山並みと海の絶景」といった身体的喜びや自然への畏敬に結びつく情景描写が多く、楽しさ・解放感や感嘆・癒しといった情動を誘発する傾向が示唆された。このように、本ケーススタディではデータ駆動型のアプローチにより、従来の人類学的手法では見えにくかった移住経験の多面的（プラットフォームごと）の様相が捉えられ、学際研究に新たな知見をもたらしている。

以上の成果から、機械学習による学術データの俯瞰的分析は、知の構造を解明する有効な手段となりうることを示された。テキストと画像双方の情報を統合的に扱うことで、学問分野間の

関係性や知識伝播のパターンをより立体的に把握できる。本研究は、学術分野の細分化と学際化が相互に作用して知識ネットワークが構築される過程を可視化し、現代の知のあり方を捉える新しいアプローチを提示した。今後、さらなる分野横断的なデータの比較分析を進めることで、より大規模なスケールで知の全体像を描き出し、学術知識の構造に関する理解を深化させることが期待される。

発表論文等

〔雑誌論文〕

- [1] 菅原裕輝, “デジタル空間の多重性”, 『文化人類学』 89(4), 581–592, 2025.