

ローカルLLMを用いた 技術動向マップの自動生成

Automatic Generation of Technology Trend Maps Using Local LLMs



中央大学 理工学術院 教授

難波 英嗣

2001年北陸先端科学技術大学院大学情報科学研究科博士後期課程修了。博士（情報科学）。東京工業大学精密工学研究所助手、広島市立大学大学院情報科学研究科准教授等を経て、2019年より中央大学理工学部教授。自然言語処理、テキストマイニングの研究に従事。

✉ nanba.11f@g.chuo-u.ac.jp TEL 03-3817-1883

1 はじめに

研究開発戦略や知的財産戦略を立案する際には、対象分野にどのような技術課題が存在し、どの技術がどの用途に結び付いているのかを把握することが不可欠である。特に、環境、エネルギー、医療、海洋安全保障のように、関連分野が広く、論文や特許が継続的に増加する領域では、個々の文献を手で読み、重要技術を抽出して整理する作業には大きな時間と費用を要する。

従来は、専門家による評価、キーワード検索、引用分析、クラスタリング、トピックモデルなどを組み合わせて技術動向を把握することが多かった。これらの手法は有効である一方、分析軸をどのように設定するか、得られたクラスターにどのような意味を与えるか、結果をどのように説明するかという点では、依然として人手への依存が大きい。また、研究開発の速度が速い分野では、一度作成した技術マップが短期間で古くなる可能性もある。

本稿では、この課題に対し、ローカル環境で動作する大規模言語モデル（Large Language Model; LLM）を用いて、技術文書集合から階層的な技術分析軸を抽出し、それに基づいて技術動向マップを生成する試みについて述べる。クラウド型 LLM や Deep Research 機能は便利であるが、入力できる文献数や情報源の制御、再現性、費用、機密情報の取り扱いに制約がある。これに対し、ローカル LLM を用いれば、分析対象文書を自ら管理し、大規模な文献集合を分割処理しながら、必要に応じて専門家が結果を確認・修正できる。

本研究の狙いは、単に文献集合を要約することではな

い。個々の論文概要から「何が課題であり、どのような手段・対象・応用が記述されているか」を抽出し、それらを階層的な分析軸として整理することで、文献単位の検索では見えにくい技術の広がりや可視化する点にある。以下では、関連研究を概観した後、「海洋の安全保障」分野を対象とした分析軸の自動構築、論文概要との対応付け、検索インターフェイス、Deep Research との比較、および今後の課題について述べる。

2 関連研究

近年、LLM を用いて大量の学術文献を整理・分類し、研究動向を把握する研究が活発に行われている。これらの研究は、文献レビュー支援、分類体系の自動構築、サーベイ論文の自動生成、技術ロードマップの作成などを対象としており、本稿で扱う技術動向マップの自動生成と密接に関連している。

Hsu らは、文献レビュー作業を支援するために、LLM を用いて関連する複数の科学論文を階層的に整理する手法 CHIME を提案している^[1]。CHIME では、各ノードがトピックのカテゴリを表し、それぞれのノードに該当する論文がリンクされたツリー構造として文献集合を整理する。また、LLM による出力結果を専門家が修正して構築したデータセットも提案している。この研究は、論文集合から階層的な構造を生成する点で本稿と共通している。一方、本稿では、文献レビュー支援そのものではなく、特定分野における技術要素を抽出し、それらを技術動向マップとして可視化・分析することを目

的としている。また、対象文献を大規模に分割処理し、ローカル LLM で処理する点も本稿の特徴である。

Wan らは、LLM を活用して非構造化テキストデータから意味のあるカテゴリラベルを自動生成し、大規模なテキスト分類を行うエンドツーエンドのフレームワーク TnT-LLM を提案している^[2]。TnT-LLM は、ユーザーの意図分析や会話ドメイン分析などのテキストマイニングにおいて、少ない人的労力で高品質な分類を実現することを目指している。この研究は、LLM を用いてカテゴリを生成し、それに基づいて大量のテキスト进行分类するという点で本稿と関連している。ただし、本稿では一般的なカテゴリ分類ではなく、技術分野の分析に利用できる多階層の分析軸を自動構築し、さらに論文概要との対応付けを行うことで、技術動向の把握や用途探索に利用する点異なる。

Kasanishi らは、自動で文献レビューを生成するタスクのための大規模データセット SciReviewGen を構築・公開している^[3]。SciReviewGen は、コンピュータサイエンス分野における 1 万件以上の文献レビューと、それらの中で引用されている 69 万件以上の論文から構成される。文献レビューの自動生成を、レビューのタイトルと章タイトルをクエリとし、引用論文のアブストラクトを入力とする、クエリ重視の複数文書要約タスクとして定式化している。この研究は、大量の論文情報を入力として、研究分野の内容を整理するという点で本稿と関連している。一方、本稿では、レビュー本文の自動生成ではなく、技術分析軸そのものを構築し、個々の論文をその分析軸に分類することを目的としている。

Zhu らは、LLM のガイドによる多角的なエンコーディングと動的クラスタリングを用いて、学術文献の階層的な分類体系を自動生成するフレームワークを提案している^[4]。また、評価のために、専門家が執筆した 156 の分類体系と 1.1 万件以上の論文からなるベンチマークデータセット TaxoBench-CS を構築している。この研究は、学術文献から階層的な分類体系を生成する点で、本稿の技術分析軸の自動構築と非常に近い。ただし、本稿では既存の分類体系に近い一般的なタクソノミーを生成するだけでなく、「海洋の安全保障」という観点を明示的に与え、特定の分析目的に沿った技術分析軸を生成する。また、生成された軸に論文概要を対応付け、検索インターフェイスを通じて利用できる形にする点に特徴がある。

Wang らは、AI のような急速に進化する分野において、LLM を用いて包括的な学術サーベイ論文を高速かつ体系的に自動生成する手法 AutoSurvey を提案している^[5]。AutoSurvey は、初期検索とアウトライン生成、サブセクションの並行ドラフト作成、統合と推敲、複数 LLM による評価と反復という手順で構成される。この研究は、LLM を用いて大量の文献を整理し、人間が理解しやすい構造にまとめるという点で本稿と関連している。一方、AutoSurvey が最終的にサーベイ論文の文章生成を目的とするのに対し、本稿では技術分析軸と論文の対応関係を構築し、利用者が技術動向を探索できるマップを生成することを目的としている。また、本稿では Deep Research 型のサービスとの比較を通じて、検索対象文献数や情報源制御の観点から、ローカル LLM による大規模処理の意義を検討している。

Yan らは、LLM を用いて高品質な学術サーベイ論文を自動生成するフレームワーク SURVEYFORGE を提案している^[6]。この研究では、人間が書いたサーベイ論文と AI が生成したサーベイ論文の間にある、アウトラインの品質や引用の正確さに関するギャップを埋めることを目的としている。また、AI による生成結果を客観的かつ多角的に評価するためのベンチマーク SurveyBench と評価指標も提案している。この研究は、LLM による文献整理や構造化の品質評価という点で本稿と関連している。本稿においても、生成された技術分析軸が妥当であるか、冗長でないか、実際の技術動向把握に有用であるかを評価することが重要であり、SURVEYFORGE のような多面的評価の考え方は今後の評価設計において参考になる。

Jansen らは、膨大な学術論文から科学的貢献を抽出し、それらがどの先行技術や知見の上に成り立っているかをリンクさせた大規模なナレッジグラフである Scientific Contribution Graph を提案している^[7]。この研究では、ACL Anthology や S2ORC から収集した約 23 万本のオープンアクセス論文から、約 200 万個の科学的貢献ノードと約 1,250 万本の前提条件エッジを抽出している。これは、論文単位の引用関係よりも細かく、文単位の短いトリプルよりも大きい粒度で科学技術の発展関係を表現するものであり、技術ロードマップの自動作成に有用なリソースである。この研究は、技術ロードマップの自動生成を目指す点で本稿と特に関

係が深い。ただし、本稿では、科学的貢献間の依存関係をグラフとして抽出するのではなく、特定分野の論文概要から技術分析軸を生成し、各軸に関連する論文を対応付けることで、技術の用途や関連領域を探索可能にする点に重点を置いている。

以上のように、既存研究では、LLM を用いた文献集合の階層化、分類体系の生成、サーベイ論文の自動生成、科学的貢献グラフの構築などが提案されている。これらはいずれも、大量の学術文献を人間が理解しやすい形に整理することを目的としており、本稿の問題意識と共通する。しかし、本稿の特徴は、第一に、海洋の安全保障という具体的な分析観点を設定し、その観点に沿った技術分析軸を自動生成する点、第二に、Web of Science から収集した 12 万件以上の論文概要を対象としてローカル LLM により大規模に処理する点、第三に、生成された分析軸に論文概要を対応付け、技術用途や関連領域を検索・確認できるインターフェイスを構築する点にある。これにより、本稿は、従来の文献レビュー支援やサーベイ自動生成とは異なり、技術動向マップの実用的な自動生成を目指すものである。

3 技術分析軸の自動構築

本稿では、海洋分野の論文概要を対象に、海洋の安全保障に関する技術分析軸を自動構築した。分析対象は、Web of Science に収録された海洋学 (Oceanography)、海洋生物学・淡水生物学 (Marine & Freshwater Biology) 分野の論文概要 122,641 件である。処理は 2 台の計算機で約 4 日を要した。各計算機は RTX A6000 (48GB) を搭載し、メモリは 64GB である。

分析軸の生成には複数のローカル LLM を比較した。候補には、gpt-oss-120b、gpt-oss-20b、Llama-3.3-70B-Instruct、Llama-3.1-8B、Mistral-7B-Instruct-v0.2、Qwen3-14B、DeepSeek-R1-0528-Qwen3-8B、gemma-3-12b-it を含めた。比較の観点は、(1) プロンプトで指定した CSV 形式を安定して守れるか、(2) 長い入力に対して応答が極端に遅くならないか、(3) 生成される分析軸が過度に一般的または冗長にならないか、(4) ローカル環境で継続的に処理できるか、である。

実験の結果、モデルサイズが大きければ常に良い結果が得られるわけではないことが分かった。大型モデルであっても、出力が指定した CSV 形式から外れたり、階層数の制約を守らなかったりする場合があった。また、長いプロンプトを与えると応答速度が極端に低下するモデルもあった。処理速度と指示追従性のバランスを考慮し、本稿の実験では gemma-3-12b-it を主に用いた。

プロンプトでは、論文概要を入力として与え、「海洋安全保障の観点から技術分析軸を構築すること」「最大 4 階層の CSV 形式で出力すること」「論文概要に記載されていない内容を追加しないこと」「表以外の出力を行わないこと」を明示した。一度に大量の論文概要を入力することはできないため、対象文書を 60 件ずつに分割し、各分割単位から一つの分析軸に統合した。その後、同一または類似の分析軸を統合し、不適切な軸や冗長な軸を削除した。

自動生成された分析軸には、海洋監視、船舶識別、違法漁業検出、密輸検出、水中センサ、AUV 検出、海洋環境データ解析など、多様な項目が含まれた。初期生成後、ChatGPT を補助的に用いて、海洋安全保障の観点から不適切または冗長な分析軸の削除を行った。その結果、分析軸数は 2,331 件から 2,235 件へと約 4% 削減された。削減率は大きくないが、ノイズとなる軸を取り除くことで、後段の検索や可視化における解釈性が向上した。

4 論文概要との対応付けと技術動向マップ

次に、生成された技術分析軸に対して、各論文概要を対応付けた。具体的には、multilingual-e5-small^[8] を用いて論文概要と分析軸をベクトル化し、両者の類似度を計算した。類似度が 0.7 以上となる場合、その論文概要を該当する分析軸に分類した。この処理により、単に分析軸の一覧を作るだけでなく、各軸がどの論文概要に基づいているのかを確認できるようにした。

この対応付けは、技術動向マップの説明可能性を高めるうえで重要である。たとえば、ある分析軸が「密輸検出」や「AUV 検出」として提示された場合、その根拠となる論文概要を確認できなければ、利用者はその妥当性を判断しにくい。提案システムでは、各分析軸に関連する論文概要を紐付け、必要に応じて「open」をクリックすると概要を表示し、「close」をクリックすると非

表示にできるようにした。これにより、分析軸と原文情報の関係を直感的に確認できる。

検索インターフェイスは二種類用意した。第一は、指定したキーワードが分析軸中出现した場合のみを表示するものである。第二は、指定したキーワードが、分析軸に分類された論文概要中出现した場合を表示するものである。前者は分析軸そのものの確認に適しており、後者は分析軸には現れていないが関連論文の本文要約に含まれる用語を手掛かりに探索する場合に有効である。検索は大文字・小文字を区別せずに実行される。

図1は、「Underwater monitoring」を検索した例である。この検索により、「Smuggling Detection（密輸検出）」「Fishing Vessel Monitoring（漁船監視）」「Autonomous Underwater Vehicle Detection（自律型水中車両検出）」「Target Tracking（ターゲット追跡）」などの分析軸が抽出された。また、船舶識別に関連して「AIS Decoding（AIS デコード）」が利用されていることも確認できた。このように、ある技術がどの用途に結び付くかを探索できる点は、研究開発テーマの発見や特許調査の出発点として有用である。

5 Deep Research との比較

本手法の特徴を明確にするため、ChatGPT、Gemini、Claude、Perplexity の Deep Research 機能を用いた分析結果と比較した。生成時間は、ChatGPT が約 10 分、Gemini が約 9 分、Claude が約 8 分、Perplexity が約 4 分であり、短時間で要約的な結果を得られる点では Deep Research が優れている。

一方で、Deep Research は、与えられたプロンプトに関連する文献を検索し、その検索結果に基づいて出力を生成する。そのため、実際に参照される文献数は多くても数十件程度にとどまる場合が多く、Claude のみ 402 件であった。海洋分野の論文 12 万件規模を分析対象とする本稿の手法と比べると、カバレッジには大きな差がある。さらに、同一の LLM を用いて同じ指示を与えても、検索される文献の違いにより、生成される分析軸が大きく変わる場合がある。

また、Deep Research では、プロンプトによって情報源を厳密に制御することが難しい。たとえば、「過去 5 年間に発表された論文と特許を情報源とする」と指定しても、条件から外れる文献や、論文・特許以外の情報源が混入する場合がある。Gemini の結果では、指示に合致した文献は 65 件中 23 件 (35%) にとどまり、

ID	axis_1	axis_2	axis_3	axis_4	2025	2024	2023	2022	2021
75	Critical Infrastructure Protection 重要インフラ保護	Shipping Lane Security 航路セキュリティ	Chokepoint Security チョークポイントセキュリティ	Autonomous Underwater Vehicles (AUVs) for Surveillance 監視用自律型水中車両 (AUV)	(20)	(29)	(54)	(59)	(38)
83	Critical Infrastructure Protection 重要インフラ保護	Shipping Lane Security 航路セキュリティ	Chokepoint Security チョークポイントセキュリティ	Strait Monitoring 海峡監視	(9)	(13)	(16)	(14)	(11)
88	Critical Infrastructure Protection 重要インフラ保護	Shipping Lane Security 航路セキュリティ	Chokepoint Security チョークポイントセキュリティ	Strategic Waterway Monitoring 戦略的水路監視	(6)	(16)	(7)	(17)	(6)
199	Maritime Cybersecurity 海事サイバーセキュリティ	Data Security データセキュリティ	AIS Data AISデータ	Position Tracking 位置追跡	(6)	(6)	(8)	(4)	(4)
228	Maritime Cybersecurity 海事サイバーセキュリティ	Data Security データセキュリティ	Operational Data 運用データ	Vessel Tracking Data 船舶追跡データ	(29)	(60)	(51)	(31)	(29)
484	Maritime Cybersecurity 海事サイバーセキュリティ	Threat Detection 脅威検出	Subsurface Threats 地下の脅威	UUV Detection UUV検出	(25)	(45)	(62)	(43)	(25)
490	Maritime Cybersecurity 海事サイバーセキュリティ	Threat Detection 脅威検出	Unmanned Systems Detection 無人システムの検出	Autonomous Underwater Vehicle Detection 自律型水中車両検出	(15)	(33)	(61)	(52)	(39)
513	Maritime Cybersecurity 海事サイバーセキュリティ	Threat Intelligence 脅威インテリジェンス	Maritime Crime Trends 海上犯罪の傾向	Smuggling Detection 密輸検出	(3)	(5)	(3)	(3)	(3)
814	Maritime Domain Awareness 海洋状況把握	Piracy 海賊行為	Piracy Tactics 海賊行為の戦術	Vessel Navigation Exploitation 船舶航行の活用	(8)	(13)	(17)	(12)	(14)
887	Maritime Domain Awareness 海洋状況把握	Sensor Technologies センサー技術	Sonar ソナー	Underwater Acoustic Surveillance 水中音響監視	(25)	(49)	(55)	(47)	(24)
945	Maritime Domain Awareness 海洋状況把握	Vessel Tracking 船舶追跡	Automated Identification 自動識別	AIS Decoding AISデコード	(14)	(24)	(12)	(12)	(6)

図1 技術分析軸を「Underwater monitoring」で検索した結果

論文・特許以外の文献が32件（49%）含まれていた。これは、情報源の透明性と再現性を重視する技術動向分析では大きな課題となる。

ローカル LLM を用いる提案手法は、Deep Research と比べて処理時間は長い。しかし、分析対象文献を利用者が明示的に指定できるため、データの範囲、時期、文献種別を管理しやすい。また、処理過程で生成された中間結果を保存できるため、分析軸の修正、再分類、しきい値の変更、別モデルとの比較などを段階的に実施できる。大量文献を対象にした再現性の高い分析という観点では、ローカル LLM による処理に利点がある。

6 考察

本研究の結果から、ローカル LLM は、大規模文献集合から技術分析軸を構築するための有力な手段となり得ることが確認できた。特に、単純なクラスタリングではクラスタの意味付けが後処理として必要になるのに対し、LLM を用いることで、最初から人間が解釈しやすいラベルを階層的に生成できる点が有用である。また、生成された軸に論文概要を対応付けることで、分析結果の根拠を確認できる。

一方で、課題も残る。第一に、LLM の出力は完全には安定しない。CSV 形式の崩れ、同義語の揺れ、抽象度の不統一、入力文書に明記されていない語の混入などが生じることがある。第二に、分析軸の統合や削除には、依然として一定の人手確認が必要である。第三に、ベクトル類似度による論文分類では、しきい値の設定が結果に大きく影響する。0.7 というしきい値は実験的に設定したものであり、対象分野や分析目的に応じた調整が必要である。

また、現在の方法では、技術間の因果関係や前提関係までは十分に扱えていない。たとえば、あるセンサ技術がどの検出技術を可能にし、さらにどの応用に結び付くのかを明示的なグラフとして表現できれば、研究開発ロードマップとしての価値はさらに高まる。今後は、分析軸と論文概要の対応付けに加え、技術、課題、対象、用途、効果の関係を抽出し、知識グラフとして整理することが重要である。

知財情報処理への応用という観点では、論文だけでなく特許文献を同時に扱うことが重要である。論文は技術シーズを早期に把握するうえで有効であり、特許は実用

化や権利化の方向性を知るうえで有効である。両者を同一の技術分析軸に対応付けることで、研究動向と出願動向の差分、学術研究から産業応用への移行、未出願領域の発見などが可能になると考えられる。

7 おわりに

本稿では、ローカル LLM を用いて大規模な論文概要集合から技術分析軸を自動構築し、それに基づいて技術動向マップを生成する方法を述べた。海洋の安全保障分野の論文概要 122,641 件を対象に、gemma-3-12b-it を用いて階層的な分析軸を生成し、重複・不適切な軸を整理した結果、2,235 項目の技術分析軸が得られた。さらに、multilingual-e5-small を用いて各論文概要を分析軸に対応付け、検索インターフェイスを構築した。

Deep Research 機能は短時間で概観を得る用途には有効であるが、参照文献数、情報源の制御、再現性の点で制約がある。これに対し、ローカル LLM を用いた方法は、処理時間を要するものの、分析対象を明示的に管理でき、大量文献に対して一貫した処理を行える。今後は、特許文献への適用、分析軸の自動統合精度の向上、技術間関係の抽出、専門家による対話的修正機能の導入を進める予定である。これにより、研究開発戦略や知財戦略の立案を支援する、説明可能で更新可能な技術動向マップの実現を目指す。

参考文献

- [1] C.-C. Hsu et al., CHIME: LLM-Assisted Hierarchical Organization of Scientific Studies for Literature Review Support, Findings of the Association for Computational Linguistics: ACL 2024, 2024.
- [2] M. Wan et al., TnT-LLM: Text Mining at Scale with Large Language Models, ACM KDD 2024, 2024.
- [3] T. Kasanishi et al., SciReviewGen: A Large-scale Dataset for Automatic Literature Review Generation, Findings of the Association for Computational Linguistics: ACL 2023, 2023.

- [4] K. Zhu et al., Context-Aware Hierarchical Taxonomy Generation for Scientific Papers via LLM-Guided Multi-Aspect Clustering, Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, 2025.
- [5] Y. Wang et al., AutoSurvey: Large Language Models Can Automatically Write Surveys, arXiv:2406.10252 [cs.IR], 2024.
- [6] X. Yan et al., SURVEYFORGE: On the Outline Heuristics, Memory-Driven Generation, and Multi-dimensional Evaluation for Automated Survey Writing, Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics, 2025.
- [7] P. A. Jansen et al., The Scientific Contribution Graph: Automated Literature-based Technological Roadmapping at Scale, arXiv:2605.15011 [cs.CL], 2026.
- [8] L. Wang et al., Multilingual E5 Text Embeddings: A Technical Report, arXiv:2402.05672 [cs.CL], 2024.