

ニーズ知識グラフ構築による異業種アナロジーの探索

朝倉卓人^{1,2,3,6} 広田航¹ 乾健太郎^{4,5,6} Benjamin Heinzerling^{5,6} Qin Dai⁵ 有馬幸介¹
¹ストックマーク ²NII ³東大 ⁴MBZUAI ⁵東北大 ⁶RIKEN
 takuto@is.s.u-tokyo.ac.jp wataru.hirota@stockmark.co.jp
 kentaro.inui@mbzuai.ac.ae benjamin.heinzerling@a.riken.jp
 daiqin@ecei.tohoku.ac.jp kosuke.arima@stockmark.co.jp

概要

大規模言語モデル (LLM) は新規アイデア生成に有用である一方、表層的連想に依存しやすく、真に意外性のある発想の創出には限界がある。本研究では、特許由来の技術シーズ知識グラフ (KG) と、ニュース記事から仮説推論により構築したニーズ KG を照合することで、異業種アナロジーを体系的に発見する枠組みを提案する。両 KG を、意味的な共通ノードを基準に対処付けることで、語彙一致に依らない意外だが妥当な対応関係を抽出する。自動車業界での評価により、本手法が LLM 単独より創造的なアイデアを創出できることを示した。

1 はじめに

革新的なアイデアは、しばしば互いに離れた領域の概念や技術を結び付けることで生まれる。近年、LLM は事業創出や研究支援などの創造的タスクに活用されつつあるが、その出力は言語的共起に基づくものが多く、真に意外性のある発想を生み出すことは依然として難しい [3, 6]。創造性において重要な意外性と現実的妥当性の両立 (セレンディピティ) の計算機的実現は、未解決の課題である。

この課題に対し、本研究では異業種アナロジーに着目する。異なる産業領域に存在する技術や課題の間に共通する機能的構造を見出すことで、言語的には近くないが背後の構造の類似性から有望と考えられる発想を導くことが可能となる。たとえば「生ビールのような泡立ちが楽しめる缶ビール」というニーズと、ある特殊な「金属塗装技術」(シーズ) が「低撥水性」というアンカー (共通ノード) で結びつけられる (図 1)¹⁾。しかし、このような対応関

係は暗黙的で疎であり、テキストの表層レベルの言語的整合性を重視する LLM のみで自動的に発見することは困難である。

そこで本研究は、明示的な意味構造に基づく探索を可能にするため、二種類の KG を構築・照合する枠組みを提案する。一つは特許文書から抽出した技術シーズ KG、もう一つはニュース記事に現れる社会的・産業的ニーズを起点とし、仮説推論により技術的機能へと展開したニーズ KG である。両者は共通の関係スキーマで表現されており、埋め込み表現を用いたアンカー (意味的に共通のノード) 対応付けにより、語彙一致に依らない意味的対応関係を同定できる。このような KG ベースのマッチングにより、LLM 単独では得にくい意外だが説明可能なシーズ・ニーズ対応を体系的に抽出できる。本研究では自動車部品領域を対象に評価を行い、提案手法が LLM による直接生成と比べて高い意外性をもつアイデアを生み出すことを示す。

2 関連研究

異なる領域間の知識対応を用いて創造性を喚起する研究は、設計工学分野におけるアナロジー設計 (Design by Analogy, DbA) として長く研究されてきた。生物模倣設計を含む古典的研究 [13, 11, 4] に加え、近年では大規模な設計データや特許コーパスを用いたデータ駆動型手法 [7] や、KG を用いて設計要素間の類似性を探索する手法 [8] が提案されている。

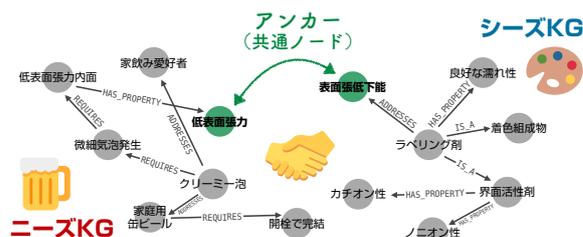


図 1 理想的なシーズ・ニーズマッチングの例

1) 本例はアサヒビール(株)とトーヨーケム(株)が協業した実例に基づく。https://www.artiencegroup.com/ja/products/metal-coatings/foam-generating-beer-can-coating.html

DbA は主に製品設計や機械設計を対象としており、特許由来の技術シーズと社会的ニーズを異業種横断で対応付ける本研究とは対象と枠組みが異なる。

近年、LLM を用いた研究 [5, 12, 10] やビジネス [6] のアイデア生成研究が活発化している。LLM は多様な分野にまたがる知識を統合し、文脈的に自然な提案を生成できる一方で、頻出概念や言語的近接性に依存しやすく、真に新規性の高い発想や異分野横断的な飛躍を生みにくいと指摘されている [3, 6]。

技術的には、テキストから要求や不満を抽出する研究 [2] や、LLM による仮説推論能力に関する研究 [1] とも関連する。本研究では、ニュース由来のニーズをそのまま用いるのではなく、仮説推論により技術的機能レベルへ展開し、特許由来の技術知識と同一の関係スキーマで表現する点に新奇性がある。

3 タスク定義：シーズ・ニーズ照合

本研究の目的は既存技術（シーズ）がもつ性質や機能と、異なる文脈で生じる社会的ニーズの間に意味的対応関係を見出し、異業種アナロジーに基づく有望なアイデア候補を抽出することである。これを二つの KG 間の照合問題として定式化する（図 2）。

入力として、二種類の KG を仮定する。一つは特許文書から抽出した技術シーズ KG であり、構成要素や特性、機能関係を含む技術的知識を表現する。もう一つはニュース記事に現れる社会的ニーズを起点とし、仮説推論によりそれらを満たすと考えられる技術的機能へ展開したニーズ KG である。重要な点として、両 KG は同一の関係スキーマ（表 1）で表現されており、構造的に比較可能である。

本タスクの出力は、シーズ KG のノードとニーズ KG のノードからなる対応ペアの集合である。これらの対応関係は、特許から抽出された技術が、当初想定されていなかった異なる領域におけるニーズを満たし得る可能性を示す。すなわち、シーズに内在する技術的性質や機能が、別文脈における未充足の要求と機能的に結び付く可能性を捉えるものである。本研究では、このような対応付けを共通スキーマをもつ二つの KG 間の照合として定式化することで、表層的な言語類似に依らず、意外性と妥当性を併せ持つ異業種アナロジーを体系的に探索する。

4 知識グラフ構築

両 KG の構築ステップで LLM を使用しているが、本研究ではいずれも GPT-4o mini を採用した。なお

表 1 シーズ・ニーズ KG 共通で用いる関係タイプ

関係	種別	意味
IS_A	構造	技術・部品とその上位概念
PART_OF	構造	構成物と部品の包含関係
HAS_PROPERTY	属性	技術や材料がもつ性質・特性
ADDRESSES	属性	技術的特徴が解決する課題
IMPROVES	属性	技術要素が向上させる性能指標
REQUIRES	属性	技術が必要とする条件や環境

具体例は付録 A に掲載する。

シーズ KG 特許文書から抽出した既存技術の知識を表現する KG である。本研究では自動車部品分野の特許 200 件を対象とし、技術要素、構成部品、性質、機能的役割などを三つ組として抽出した。用いる関係タイプ（表 1）は、部品階層や技術的特性を明示的に表現できるよう設計されている。知識抽出には LLM を用い、出力形式を制約するプロンプトにより、関係タイプの一貫性とノイズ低減を図った。エンティティ表現は正規化を行わず自由記述とし、粒度を保つことで後述するニーズ KG との意味的比較を可能にしている。その結果、語彙的な一致はほとんど生じず、後段の対応付けは意味的類似性に依存する設計となっている。

ニーズ KG ニュース記事に現れる社会的・産業的ニーズを起点として構築される。ただし、ニュース記事自体には技術的詳細が含まれない場合が多いため、本研究では二段階の仮説推論を用いてニーズを技術的機能表現へと展開する。まず、ニュース記事から製品や産業に関する不満・要望を LLM で抽出し、産業製品に該当しない事例を除外する。その後、各ニーズに対して「どのような技術的方向性が考えられるか」をやはり LLM を用いて推定し（第一段階）、さらにそれをシーズ KG と共通のスキーマ（表 1）を用いた部品・特性・機能レベルの表現へと LLM で具体化（第二段階）する。

各 KG の特徴 両 KG の基本的な統計量を表 2 に示す。シーズ KG は比較的コンパクトである一方、技術的属性や機能関係を豊富に含む。一方、ニーズ KG は大規模であるが、各ニーズに対応する部分グラフは小さく、構造も浅い。両者は規模や深さに違いがあるものの、これはデータ特性に由来するものであり、照合自体を妨げるものではない。

5 シーズとニーズの対応付け

ここでは前節で構築した二つの KG を照合し、対応候補を抽出・順位付けする方法を述べる。

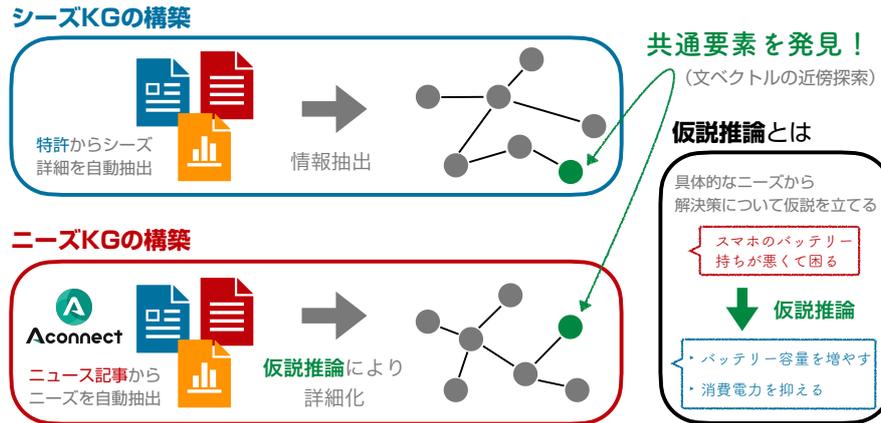


図2 技術シーズと社会的ニーズの対応付け

表2 各 KG の基本統計量

統計量	シーズ KG	ニーズ KG
総トリプル数	16,270	912,386
平均トリプル数	81.35	9.21
標準偏差	35.03	1.63
中央値	76	9
平均トリプル数	4.15	2.22
標準偏差	1.20	0.54
中央値	4	2

5.1 アンカーに基づく意味的対応付け

シーズとニーズの対応付けは、両 KG に含まれるノード間の意味的類似性に基づいて行う。本研究では、各ノードのテキスト表現を日本語 Sentence LUKE²⁾を用いて文(句)レベルの埋め込みに変換し、シーズ KG とニーズ KG の全ノード対についてコサイン類似度を計算する。類似度が所定の閾値を超えるノード対を、両 KG を接続するアンカー(意味的な共通ノード)として採用する。

シーズ KG およびニーズ KG はいずれも自由記述のエンティティ表現を用いており、語彙的一致はほとんど生じない。そのため、本手法では文字列一致や表層的な類似に依らず、埋め込みに基づく意味的近接性のみを用いて対応付けを行う。アンカーが同定されると、そのノードが属する特許由来のシーズと、ニュース由来のニーズが一つの対応ペアとして結び付けられる。

5.2 シーズ・ニーズペアのランキング

得られるシーズ・ニーズ対応ペアは多数に及ぶため、創造的価値の高い候補を優先して提示する順位

2) <https://huggingface.co/sonoisa/sentence-luke-japanese-base-lite>

付けを行う。三種類のランキング戦略を比較する。

まず、ランダム順位付けは、対応ペアを無作為に並べたものであり、ベースラインとして用いる。次に、類似度順位付けは、アンカーとなるノード対の埋め込み類似度が高い順に対応ペアを並べる方法であり、意味的に近い対応を優先するため妥当性は高いが、意外性は低くなる傾向がある。最後に、希少度順位付けでは、アンカーノードの出現頻度に基づき、出現頻度の低い概念を含む対応を高く評価する。希少な概念は限定的な文脈でのみ用いられるため、異分野間の橋渡しとして機能しやすく、意外性の高い対応を生みやすいと考えられる。

6 実験と評価

本研究では、自動車部品分野を対象としてシーズ・ニーズ対応付け手法を評価した。評価対象として、4つの方法を比較する。LはLLMによりニーズから直接アイデアを生成するベースラインである。Bはアンカーに基づくシーズ・ニーズ対応付けを行い、対応ペアをランダムに提示する方法である。Sは同様の対応付けを行った上で、アンカーノード間の意味的類似度に基づいて順位付けする方法である。Rはアンカーノードの出現頻度に基づく希少度を用いて順位付けする方法である。各方法により生成された対応ペアまたはアイデアを混合・匿名化した上で、人手評価者4名およびLLM-as-a-Judgeにより、実現性、影響度、意外性、妥当性の4観点から評価を行った。

6.1 評価結果とアノテーション一致率

評価結果を表3に示す。人手評価の結果、LLMによる直接生成(L)は実現性および影響度において

表3 人間アノテータと LLM-as-a-Judge の評価結果

手法	人間アノテータ				LLM-as-a-Judge			
	実現性	影響度	意外性	妥当性	実現性	影響度	意外性	妥当性
KG / ランダム (B)	0.47	0.81	1.21	0.46	1.10	1.34	1.45	1.26
KG / 類似度 (S)	0.46	0.86	1.28	0.50	1.34	1.26	1.43	1.54
KG / 希少度 (R)	0.75	0.93	1.07	0.69	1.28	1.30	1.50	1.46
LLM (L)	0.88	1.31	0.86	0.84	1.68	1.75	1.05	1.66

表4 アノテータ間一致率 (Krippendorff の α ; ordinal)

グループ	実現性	影響度	意外性	妥当性
人間アノテータ ($N = 4$)	0.313	0.282	0.328	0.231
LLM-as-a-Judge ($N = 5$)	0.786	0.714	0.628	0.638
両方 ($N = 9$)	0.239	0.360	0.341	0.176

最も高い評価を得た一方、意外性の評価は相対的に低かった。これに対し、KG に基づく提案手法 (B, S, R) は、いずれも LLM ベースラインより高い意外性を示し、異分野間のより距離のある対応関係を抽出できていることが確認された。特に方法 R は、方法 B や S と比べて、意外性を保ちつつ実現性や妥当性も比較的高く、全体としてバランスの取れた結果を示した。LLM-as-a-Judge においても、絶対値は高めに評価される傾向が見られたものの、方法間の相対的な傾向は人手評価と概ね一致していた。

アノテータ間一致率を Krippendorff の α [9] により評価した結果 (表 4)、人手評価者間の一致率はいずれの評価軸においても低く、意外性や妥当性といった創造性に関わる指標の主観性が改めて確認された。一方、LLM を評価者として用いた場合には高い自己一致性が得られたが、人手評価者との間では判断基準に差異が見られた。これらの結果は、異業種アナロジー探索のような創造的タスクの評価が本質的に主観的で、定量評価が難しいことを示唆している。一方で、低い一致率にもかかわらず、手法ごとの平均的傾向はアノテータ・LLM-as-a-Judge の違いによらず一貫することは注目に値する。

6.2 事例研究

定性分析として各手法で高評価を得たシーズ・ニーズ対応を比較したところ、明確な傾向の違いが確認された (具体例は付録 B)。LLM による直接生成 (L) は、ガスセンサと大気モニタリング、特定ガス検出と環境管理といった、同一または近接分野内での対応を多く生成し、実現可能性や妥当性は高い一方で意外性は低かった。ランダム順位付け (B) は、酸素イオン伝導層と培養肉における酸素供給の

ように、工業材料技術とバイオ生産の要求を結び付ける大胆な対応を多く含み、高い意外性を示したが、妥当性や実装可能性の評価はばらついた。類似度順位付け (S) は、膜材料と水再利用、保護膜と表面劣化対策など、語義的・機能的に近い概念同士の対応が中心となり、安定したが保守的な結果となった。希少度順位付け (R) は、圧力損失、熱膨張係数、温度分布の偏りといった抽象度の高い物理・機能概念を媒介として異分野を接続する対応が多く、意外性と妥当性の両立が最もよく達成されていた。

一方、LLM-as-a-Judge による評価では、手法間の差異は大きく弱まり、ほとんどの対応が高い実現可能性・影響度・妥当性として評価された。意外性の分布も狭く、酸素イオン伝導層と培養肉の酸素供給のような異分野横断的対応と、ガスセンサと環境計測のような分野内対応との差が十分に反映されなかった。この結果は、LLM による評価が文脈的一貫性や表層的妥当性を重視し、構造的な異分野性や機能的飛躍を過小評価する可能性を示している。

7 結論

本研究では、特許とニュース記事という出自の異なる二つの KG を照合することで、創造性の高いアイデア候補を提示する手法を提案した。評価の結果、LLM による直接生成は実現可能性や影響度に優れる一方で意外性に乏しく、KG に基づく対応付けはより意外性の高い対応を多く生み出すことが確認された。特に、希少な性質を媒介とした対応は、異分野間の機能的類似性を捉えつつ、比較的高い妥当性を示した。

これらの結果は、LLM と KG が創造的支援において補完的な役割を果たし得ることを示している。LLM は分野内で整合的な発想に強い一方、KG に基づく構造的探索は、言語的近接性に依らない意味的飛躍を可能にする。評価実験からは、人手評価と LLM 評価の間に差異が存在することも示された。今後は、より豊かな構造情報を用いた対応付けや、創造性評価手法の検討を進める必要がある。

8 謝辞

アノテーションや技術的議論に協力してくださったストックマークの高橋洸丞氏と進藤尚希氏に感謝する。

参考文献

- [1] Balepur, N., Ravichander, A., Rudinger, R.: Artifacts or abduction: How do LLMs Answer Multiple-choice Questions Without the Question?. In: Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics. (2024)
- [2] Cai, H., Xia, R., Yu, J.: Aspect-category-opinion-sentiment Quadruple Extraction With Implicit Aspects and Opinions. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. (2021)
- [3] Chen, J. et al.: GenQA: Generating Millions of Instructions from a Handful of Prompts. arXiv:2406.10323. (2024)
- [4] Fu, K., Moreno, D., Yang, M., Wood, K. L: Bio-inspired Design: An Overview Investigating Open Questions From the Broader Field of Design-by-Analogy. *Journal of Mechanical Design*, 136(11). (2014)
- [5] Guo, S., et al.: IdeaBench: Benchmarking Large Language Models for Research Idea Generation. In: Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2, pp. 5888–5899. (2024)
- [6] Hirota, W., et al: Overview of PBIG Shared Task at AgentScen 2025: Product Business Idea Generation from Patents. In: Proceedings of the 2nd Workshop on Agent AI for Scenario Planning. (2025)
- [7] Jiang, S., Hu, J., Wood, K. L., Luo, J.: Data-driven Design-by-Analogy: State-of-the-art and Future Directions. *Journal of Mechanical Design*, 144(2). (2022)
- [8] Jing, L., Huang, M., Li, Q., Dou, Y., Feng, D., Jiang, S.: Knowledge Graph-assisted Design-by-Analogy: Promoting Product Innovation Through Structured Analogical Knowledge Retrieval. *AI EDAM*, 39. (2025)
- [9] Krippendorff, K.: *Content Analysis: An Introduction to Its Methodology*. 4th edn. Sage Publications (2018).
- [10] Ueda, K., et al. Exploring the Design of Multi-agent LLM Dialogues for Research Ideation. In: Proceedings of the 26th Annual Meeting of the Special Interest Group on Discourse and Dialogue, pp. 322–337. (2025)
- [11] Shu, L.H., Ueda, K., Chiu, I., Cheong, H.: Biologically Inspired Design. *CIRP annals*, 60(2), pp. 673–693. (2011)
- [12] Si, C., Yang, D., Hashimoto, T.: Can LLMs Generate Novel Research Ideas? A Large-scale Human Study with 100+ NLP Researchers, arXiv:2409.04109. (2024)
- [13] Ullman, D. G.: *The Mechanical Design Process*, McGraw-Hill, New York. (1992)

A シーズ・ニーズ KG の具体例

以下に、構築したシーズおよびニーズ KG の一部を示す。

シーズ KG の一部			ニーズ KG の一部		
Subject	Relation	Object	Subject	Relation	Object
光検出器	IS_A	焦電型光検出器	ポリマー配合材	IS_A	新素材
光検出器	PART_OF	光検出装置	ポリマー配合材	HAS_PROPERTY	多様な表面に適応する接着性
焦電素子	PART_OF	光検出器	ポリマー配合材	ADDRESSES	吸盤接着の制約
光検出器	HAS_PROPERTY	高い電気特性	ポリマー配合材	PART_OF	添加剤
光検出器	ADDRESSES	誤動作の発生			

B 各手法の出力例

LLM ベースラインと KG ベースの提案手法の各バリエーションの出力のうち、人間アノテータによる評価 (全項目の平均) で高く評価されたシーズ・ニーズのマッチング事例トップ 3 件を以下に示す。なお、シーズやニーズの内容は直感的に理解しやすいよう要約・編集してある。

手法	ID	ニーズ (要約)	シーズ (要約)	共通ノード	平均評価
LLM	L0078	大気中の有害 NO をリアルタイムで検出したい	NO 濃度を測定するガスセンサ技術	NO 濃度 ↔ NO 濃度	U=0.25, F=2.00, I=1.50, P=2.00
	L0071	温室環境で特定ガスを高精度に検知したい	保護カバー付きガスセンサ	ガスセンサ ↔ 特定ガス検出	U=0.25, F=1.75, I=1.50, P=2.00
	L0028	医薬品粉体の検査精度を向上させたい	ハニカムフィルタ検査装置	検査精度向上 ↔ 粒子分析	U=1.00, F=1.00, I=2.00, P=1.00
KG (B)	B0014	培養肉で酸素不足による組織壊死を防ぎたい	電解セル/セルスタック装置	酸素イオン伝導 ↔ 酸素透過材料	U=2.00, F=0.75, I=1.75, P=0.75
	B0042	固体電解質が膨張で亀裂を生じる問題	ガスセンサの製造・制御技術	多層固体電解質 ↔ 多層構造	U=1.25, F=1.00, I=1.50, P=0.75
	B0095	産業用途でのガス供給不足や供給管理の課題	ガスセンサ装置	出口側ガス流路 ↔ ガス供給工程	U=0.25, F=1.25, I=2.00, P=1.00
KG (S)	S0018	農業分野で再生水の利用が進んでいない	半導体用膜材料	膜 ↔ 膜	U=1.25, F=1.00, I=1.75, P=1.25
	S0002	染毛時に保護成分が失われ髪が損傷する	ガスセンサ構造	保護膜 ↔ 保護膜	U=2.00, F=0.75, I=1.00, P=0.50
	S0010	ナンバープレートの視認性がカバーで低下する	蛍光体/照明デバイス	反射膜 ↔ 反射膜	U=1.00, F=1.25, I=0.50, P=1.25
KG (R)	R0008	ナノファイバフィルタで圧力損失が発生する	多孔質セラミックフィルタ	圧力損失層 ↔ 同一現象	U=0.50, F=1.50, I=1.50, P=1.50
	R0021	液体水素貯蔵で水素漏洩を防ぎたい	ハニカム担体評価装置	温度分布偏り ↔ 断熱特性	U=2.00, F=0.75, I=1.25, P=0.75
	R0010	熱膨張差でペロプスカイト太陽電池が劣化する	触媒担体基板の製造技術	熱膨張係数 ↔ 材料種別	U=1.25, F=1.00, I=1.25, P=0.75