

GraphRAG とローカル LLM を用いた産学連携マッチング可視化のためのナレッジグラフの提案

奥村 拓喜† 佐野 睦夫†

概要: 本稿では、産学連携における「企業のニーズ」と「大学研究室のシーズ」間の語彙的・意味的ギャップを解消し、かつマッチングの根拠を説明可能な形で提示する支援システムを提案する。従来のマッチング手法は、キーワードの一致や単純なベクトル検索に依存しており、用語の不一致による取りこぼしや、推奨理由のブラックボックス化が課題であった。本研究では、GraphRAG (Retrieval Augmented Generation) のアプローチを拡張し、大規模言語モデル (LLM) を用いて企業の課題解決に有効なマッチングを提案し、グラフに示すシステムを開発した。実在企業と研究室のデータをマッチングし、本システムが従来の単語一致では発見困難な潜在的な連携候補を提示できることを示す。実験では、マッチシステム及び入力データから意味的に離れた単語同士でも繋げられる類推システムが問題なく動作していることがわかった。現状、関連単語が有効なものかを判断するデータや根拠が少ないため、データを変えた検証や、回数を重ねることが必要である。

1. はじめに

1.1 研究背景

近年、オープンイノベーションの重要性が高まる中、大学が保有する専門知（シーズ）と企業が抱える課題（ニーズ）を効果的に結びつける産学連携の仕組みが求められている。しかし、中小企業などが作成する課題記述書は一般的なビジネス用語で書かれることが多く、一方で大学の研究室情報は高度に専門的な学術用語で記述される傾向がある。この「語彙のギャップ」により、単純なキーワード検索では適切なマッチングが行われないう問題が存在する。また、近年の LLM を用いた推薦システムは高い精度を誇る一方で、なぜその研究室が選ばれたのかという「根拠」が不明確になりがちであり、ステークホルダーの納得感を得にくいという課題（ハルシネーションのリスク含む）がある。

また、単純に GraphRAG を使用したマッチシステムでは、単語の意味ベクトルなどを用いたマッチを行うが、「教育」と「研修」のような、本来の意味合いは近いものがマッチしづらいという問題もある。

1.2 研究目的

本研究の目的は、企業と研究室間での語彙のギャップを埋めつつ、マッチングの論理的根拠を可視化できる意思決定支援システムを構築することである。具体的には、以下の技術的アプローチにより課題解決を図る。

- 知識グラフの構築: 企業資料と研究室情報を構造化データとして統合する。
- LLM による概念拡張 (Synergy Inference): 企業の抽象的な課題から、解決策となる具体的な「技術概念」を LLM に推論させ、知識グラフ上の新たなノードとして動的に追加する。
- マッチング過程の可視化: 企業と研究室がどのような概念やキーワードを通じて結びついたかを、グラフ上の経路 (パス) として可視化する。

本システムを通じて、ソイチャレにおける産学連携の効率化と成功率向上に貢献することを目指す。

2. 関連研究

2.1 企業と研究機関のマッチングに関する先行研究

産学連携におけるシーズとニーズのマッチングは、古くから重要な研究領域とされてきた。具体的には、特許や論文のテキストマイニングによる可視化 (山本[1], 堀江ら[3]) や、当時の政策を反映した体系化の実践 (谷口[2]), あるいはオープン・イノベーションにおけるマッチングプロセス (内海[4]) やその課題点 (澤田[5]) に関する議論などが挙げられる。しかし、これら先行研究の大半は 10 年以上前のものであり、近年の産学連携の変容が考慮されていない。さらに、本研究で着目する Graph RAG を用いたマッチングシステムに関する報告は見当たらない。以上の背景から、本研究は「最新の産学連携環境への対応」と「Graph RAG の新規適用」という独自性を有するものである。

2.2 Graph RAG (Retrieval Augmented Generation) に関する研究

LLM の信頼性向上に寄与する Graph RAG に関しては、主に海外で研究が進められている。具体的には、KG-RAG4SM によるハルシネーション対策 (Chuangtao ら [6])、専門分野におけるグラフ検索の検証 (Qinggang ら [7])、あるいは教育現場での意思決定支援 (Liu ら [8]) などが挙げられる。しかし、これらの先行研究は国外のものが大半を占めており、国内での蓄積は乏しい。さらに、産学連携を主題とした Graph RAG の活用事例は世界的にも前例が乏しいため、本研究はこの未開拓領域に取り組む点において新規性を有する。

2.3 筆者自身の先行研究

この研究は、筆者 [9] の「生成 AI を用いた企業のニーズと研究室の持つシーズのマッチングシステム」の論文から引き続き研究を行っているものである。

3. システムの詳細

3.1 システム概要

本研究では、企業ニーズと研究シーズの架け橋となるマッチングシステムを提案する。本システムは、(1)企業・研究室情報の構造化入力、(2)RAG エンジンによるマッチング実行、(3)知識グラフによる根拠の可視化、という3つの主要モジュールから構成される。実装にあたっては、チャットボットや PDF 解析といったインターフェース機能の充実よりも、Graph RAG を用いたマッチングロジックの確立を優先課題とした。そのため、現在は既存資料から抽出された情報の処理に特化した設計となっている。また、本システムはローカル環境によって動作する。

3.2 システムの流れ

本システムの流れについて、詳細に示す。

1. 情報収集と正規化 企業の PDF・Web・ヒアリング情報と、研究室のシーズ情報を収集し、システム共通の JSON フォーマットに統一 (正規化) して格納する。このセクションは使用者が手動で行う。
2. 意味情報の拡張 (類推) 企業キーワードから意味的に類似する単語を類推・生成する。これらを新たなノードとして定義することで、表記の違いを超えた柔軟なマッチングを可能にする。
3. RAG による推論 構築されたデータ構造上で Graph RAG を実行し、ニーズとシーズのマッチングを行う。同時に、ノード間の接続経路を解析し、マッチングの論理的な

繋がりを提示する。

4. グラフ描画 マッチング結果を構成する企業・研究室・キーワードをノードとしたグラフを生成・出力し、関連性の構造を可視化する。

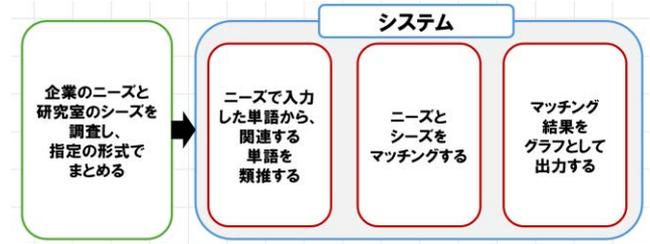


図1 システムの流れ

システムの大まかな流れを図示したものを、図1に示す。

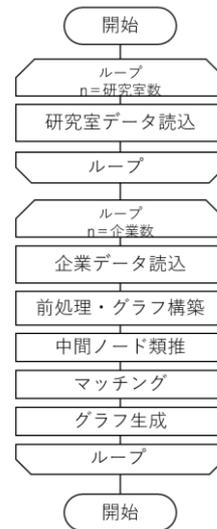


図2 フローチャート

システム内のフローチャートは、図2のようになる。

3.3 システム利用に伴う前準備

企業情報: システムを利用する企業の概要をまとめた JSON ファイル。これには企業のニーズ、ニーズを単語に言いかえたもの、企業の特徴、製品概要などが含まれる。現状、この中身の情報は企業関係者からヒアリングした内容などを筆者の自己判断でまとめ、指定の形式で記述する方式を取っている。

研究室情報: 大阪工業大学内の各研究室の概要をまとめた JSON ファイル。これには、研究室の専門分野、過去の研究実績、得意な技術、保有する設備、教員の専門性などのシーズ情報が含まれる。現状、この中身の情報は大学の研究室紹介および、研究室の HP の内容を筆者の自己判断でまとめ、指定の形式で記述する方法を取っている。

4. 実験

本実験は大学での社会課題解決の取り組みにおける企業ニーズと研究室ニーズのマッチングを目標として行う、

4.1 実験環境

使用したデバイスは大学で使用しているデスクトップPC、開発環境、及び実行環境はvscodeにてオフライン下で行った。ローカルLLMはOllmaを使用している。

4.2 使用した資料

社会課題解決の取り組みの管理をを担当している教員から、2025年度の取り組みに参加している自治体と対応研究室的の表を提供していただいたため、それを使用する。また、企業の情報は、取り組みに参加している企業2社から直接ヒアリングした内容を筆者がまとめたものを使用している。研究室データに関しては、取り組みにおいて上記2社と関連の深い活動を行っていた3つの研究室を対象とした。

よって今回の実験は、企業データ2つ、研究室データ3つで行った。

企業データは、以下のような構成である。

```
{"id": "c_0001", "name": "(企業名)", "keywords": ["技術継承", "指導方法の確立", "マーケティング"], "goal": "若手への有効な技術継承方法の確立", "budget": "mid", "prefecture": "大阪府", "ip_policy": "joint", "equipment_required": ["(企業の特徴)"]}
```

研究室データは、以下のような構成である。

```
{"id": "l_0103", "univ": "(大学名)", "pi": "(教授名)", "location": {"prefecture": "大阪府"}, "keywords": ["認知障害者のリハビリテーション", "人間と心が通じ合えるコミュニケーションAI", "産学連携システム", "人工知能", "深層学習による災害状況推定"], "records": {"papers": [{"title": "大規模マルチモーダルモデルを用いた行動理解による認知リハビリテーション振り返り支援システム"}, {"title": "eVTOL社会受容性学修のためのインタラクティブなVRシミュレータ"}]}, "strengths": "行動理解や感情理解を通じて課題を解決する"}
```

4.3 実験内容

上記の資料をシステムに入れ、3.2に示した手順を行い、実際にマッチング、及びグラフの可視化ができるかどうかを実験する。また、推論で出力された単語が妥当か同課についても検討する。

4.4 実験データ

実験に使用したデータの概要を明示する。

情報保護の観点で、社名や教授名等一部内容は伏せている。

研究室データ

```
{"id": "l_0101", "pi": "研究者 a", "location": {"prefecture": "大阪府"}, "keywords": ["マルチエージェント向けモデリング&シミュレーション", "自律分散協調", "意思決定支援", "人工知能", "最適化(ヒューリスティクス)"], "records": {"papers": [{"title": "社会システムの解析・設計・制御や、有事(災害時等)の際の意思決定に資する研究"}, {"title": "実世界の社会システム(交通等)を高精度に模擬するためのモデリング&シミュレーションに関する研究"}]}, "strengths": "個人や偉人のAIデータベース化へ関心"}
```

```
{"id": "l_0102", "pi": "研究者 b", "location": {"prefecture": "大阪府"}, "keywords": ["認知リハビリテーション", "教育支援アプリケーション", "エンターテインメント", "行動分析", "最適化(ヒューリスティクス)"], "records": {"papers": [{"title": "漫画背景自動生成システム"}, {"title": "人の視線情報から無意識化の興味を取得する"}]}, "strengths": "人の生活などからどのような課題があるのかを調査し、その課題に対して取り組む"}
```

```
{"id": "l_0103", "pi": "研究者 c", "location": {"prefecture": "大阪府"}, "keywords": ["認知障害者のリハビリテーション", "人間と心が通じ合えるコミュニケーションAI", "産学連携システム", "人工知能", "深層学習による災害状況推定"], "records": {"papers": [{"title": "大規模マルチモーダルモデルを用いた行動理解による認知リハビリテーション振り返り支援システム"}, {"title": "eVTOL社会受容性学修のためのインタラクティブなVRシミュレータ"}]}, "strengths": "行動理解や感情理解を通じて課題を解決する"}
```

企業データ

```
{"id": "c_0001", "name": "A社", "keywords": ["技術継承", "指導方法の確立", "マーケティング"], "goal": "若手への有効な技術継承方法の確立", "budget": "mid", "prefecture": "大阪府", "ip_policy": "joint", "equipment_required": ["(企業の特徴)"]}
```

```
{"id": "c_0002", "name": "テイクシステムズ", "keywords": ["日報自動作成", "若手が相談しやすい環境"], "goal": "仕事内容やにかけた時間を自動で測定し記録するシステム", "budget": "mid", "prefecture": "大阪府", "ip_policy": "joint", "equipment_required": ["検査機器", "検査の自動化"]}
```

4.5 実験結果

4.3 の内容に基づいて、実験を行った結果を以下に示す。

4.5.1 1 回目

1 回目の実験のみ、システムの動作確認のため、研究室データを仮の物で行っている。その他条件は同じである。マッチング結果は及び出力画像は以下ようになった。情報保護の観点から、実験結果に支障のない箇所を後付けで変更している。

=== A 社 候補 TOP3 ===

- 〇〇大学/山田太郎: score=0.132 (graph=0.045, sim=0.917)
 - A 社 / 若手への有効な技術継承方法の確立 → 技術継承 → VR 研修 → 〇〇大学/山田太郎 [0.007]
 - A 社 / 若手への有効な技術継承方法の確立 → 技術継承 → 作業標準化 → 〇〇大学/山田太郎 [0.005]
- △△大学/佐藤花子: score=0.125 (graph=0.043, sim=0.860)
 - A 社 / 若手への有効な技術継承方法の確立 → 技術継承 → 引張試験 → △△大学/佐藤花子 [0.005]
- 大学/田中一郎: score=0.123 (graph=0.040, sim=0.869)

=== B 社 候補 TOP3 ===

- 〇〇大学/山田太郎: score=0.125 (graph=0.043, sim=0.863)
 - B 社 / 仕事内容やにかけた時間を自動で測定し記録するシステム → 日報自動作成 → 作業標準化 → 〇〇大学/山田太郎 [0.005]
 - B 社 / 仕事内容やにかけた時間を自動で測定し記録するシステム → 若手が相談しやすい環境 → VR 研修 → 〇〇大学/山田太郎 [0.005]
- 大学/田中一郎: score=0.124 (graph=0.043, sim=0.854)
 - B 社 / 仕事内容やにかけた時間を自動で測定し記録するシステム → 日報自動作成 → 画像処理 → □□大学/田中一郎 [0.005]
- △△大学/佐藤花子: score=0.124 (graph=0.043, sim=0.854)
 - B 社 / 仕事内容やにかけた時間を自動で測定し記録するシステム → 日報自動作成 → 破断面観察 → △△大学/佐藤花子 [0.005]

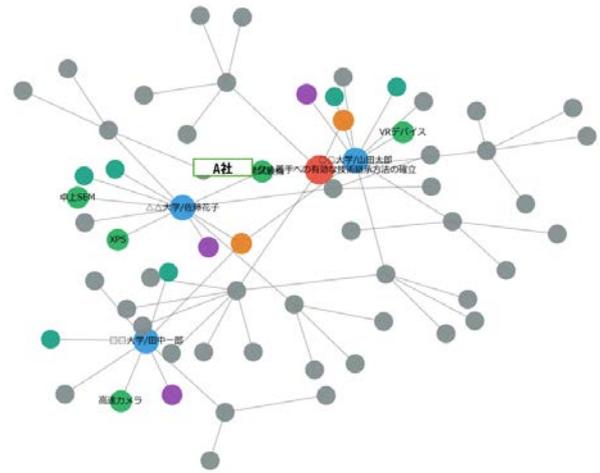


図3 実験1 A社全体グラフ

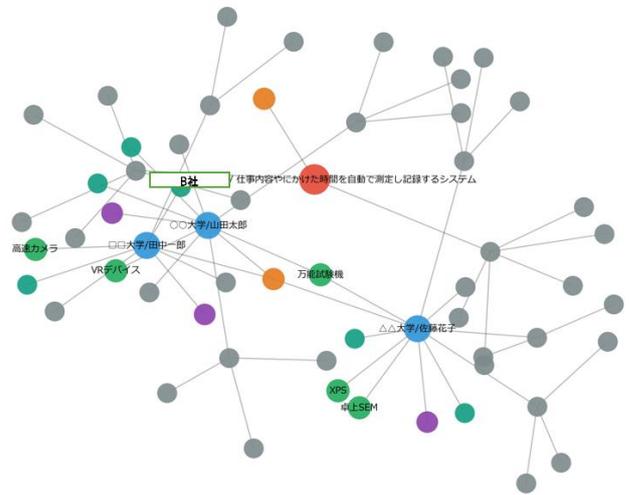


図4 実験1 B社全体グラフ

図3, 図4は出力されたグラフの一部である。全体を移したものであり、見づらいが実験結果のマッチ度が高い順に、会社のノード(赤色)と研究室のノード(青色)が高くなっていることがわかる。これとは別に1社と1研究室にフォーカスしたものもあるが、ノード同士が繋がっていないなど、実用に足るものではないので省略した。

またマッチング結果からも、元のデータにある単語のみを使用しており、3.2で示したようにLLMによる推論が反映されていない。

4.5.2 2 回目

データの重みやコード内容を調整し、LLMによる推論をより強く反映するようにした実験2の結果の一部が以下ようになった。

情報保護の観点から、実験結果に支障のない箇所を後付けで変更している。

=== A 社 候補 TOP3 ===

[Synergy] 新規ノード (2):

id=concept:技術継承の指導方法, type=Concept, label=技術継承の指導方法

id=kw:技術継承の指導方策, type=Keyword, label=技術継承の指導方策

1. 大阪工業大学 / 研究者 b : score=0.200 (graph=0.124, sim=0.885)

A 社 / 若手への有効な技術継承方法の確立 → マーケティング → エンターテインメント → (大学名) / 研究者 b [0.177]

A 社 / 若手への有効な技術継承方法の確立 → 指導方法の確立 → 教育支援アプリケーション → (大学名) / 研究者 b [0.176]

2. 大阪工業大学 / 研究者 c : score=0.199 (graph=0.123, sim=0.885)

A 社 / 若手への有効な技術継承方法の確立 → lab:l_0101 → 人工知能 → (大学名) / 研究者 c [0.176]

3. 大阪工業大学 / 研究者 a : score=0.198 (graph=0.123, sim=0.872)

A 社 / 若手への有効な技術継承方法の確立 → マーケティング → 最適化(ヒューリスティクス) → (大学名) / 研究者 a [0.176]

この結果は、A 社と他研究室とのマッチング結果を示したものである。「技術継承の指導方法」に対して、「技術継承の指導方策」という単語が LLM によって推定できていることがわかる。

しかし、lab:l_0101 など、本来マッチに入らない研究室の ID がノードとして使用されている、グラフが結果と同じにならないなどの不具合が起きている。そのため本実験でのグラフは省略している。

4.5.3 3 回目

2 回目のような類推では、ほぼ意味が変わらず推定の意味が無いと判断した為、LLM の推定のプロンプトや、不具合の修正などを行い、実験 3 を行った。結果の一部を以下に示す。情報保護の観点から、実験結果に支障のない箇所を後付けで変更している。

=== A 社 候補 TOP3 ===

[Synergy] Generated Concepts: ['教育学', '人材育成理論', '組織変革理論', 'キャリアデザイン', '組織文化構築']

1. 大阪工業大学 / 研究者 a : score=0.286 (graph=0.221, sim=0.872)

マッチ経路:

A 社 / 若手への有効な技術継承方法の確立 → 教育学 → 人工知能 → (大学名) / 研究者 a [0.304] (Synergy)

A 社 / 若手への有効な技術継承方法の確立 → 組織文化構築 → 意思決定支援 → (大学名) / 研究者 a [0.261] (Synergy)

2. 大阪工業大学 / 研究者 b : score=0.286 (graph=0.219, sim=0.885)

マッチ経路:

A 社 / 若手への有効な技術継承方法の確立 → 教育学 → 教育支援アプリケーション → (大学名) / 研究者 b [0.313] (Synergy)

A 社 / 若手への有効な技術継承方法の確立 → 組織文化構築 → 行動分析 → (大学名) / 研究者 b [0.261] (Synergy)

3. 大阪工業大学 / 研究者 c : score=0.280 (graph=0.213, sim=0.885)

マッチ経路:

A 社 / 若手への有効な技術継承方法の確立 → キャリアデザイン → 人間と心が通じ合えるコミュニケーション AI → (大学名) / 研究者 c [0.260] (Synergy)

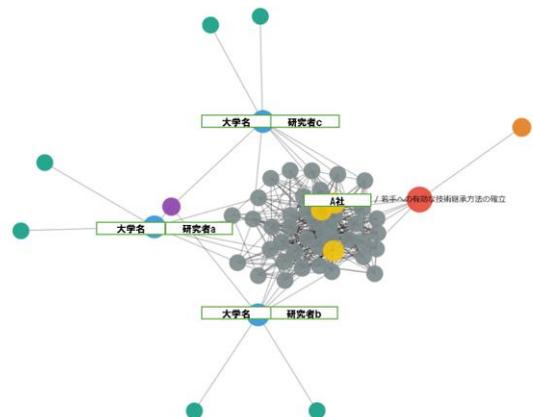


図5 実験3 A社全体グラフ

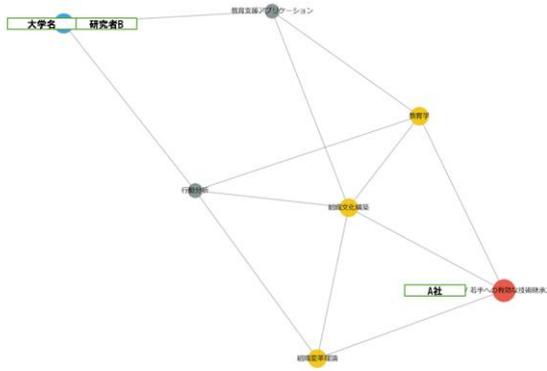


図6 実験 A社-研究者b個別グラフ

実験3では、「技術継承の指導方法」に対して、[教育学, '人材育成理論', '組織変革理論', 'キャリアデザイン', '組織文化構築']という単語が推定できており、マッチングにも使用されている。図6からみても、それらのノード（黄色）を経てマッチングをしていることがわかる。

また、実験1と同じ方式でグラフを出すと、図5のようになる。今回の結果では、3研究室ともマッチ度の差が警備の為、距離の違いは見られない。

5. 考察

実験3の結果から、入力された情報からローカルLLMにより近い意味の単語を推定し、マッチングに使用することはできていると考える。また、「若手への有効な技術継承方法の確立 → 教育学 → 教育支援アプリケーション」のように、それらしい経路を通りマッチできている。出力されたグラフからも、類推単語からの経路でマッチしていることがわかる。しかし、図5のように、各研究室の差が見受けられずらく、全体グラフの意味合いが減少しているのは課題点と言える。また現状では検証量が少なく、結果の妥当性の判断などが難しい。

6. おわりに

本稿では、GraphRAGとLLMの推論能力を融合させた産学連携マッチングシステムを提案した。

企業の課題から解決策となる概念を動的に生成し、知識グ

ラフ上で可視化することで、語彙のギャップを超えたマッチングと説明性の両立を目指したが、スコアリングの方法、出力結果の妥当性の評価の観点で問題がある。今後は、スコアリング精度の向上や、より多くのデータや試行回数による結果の分析、別の研究室も含めた結果の定量的な評価に取り組みたい。

参考文献

- [1] 山本 外茂男,産学連携のマッチング性分析におけるテキストマイニングの有効性, 情報の科学と技術, 59 巻 6 号 p. 291-297.2009
- [2] 谷口 邦彦, 産学官連携によるイノベーション創出プログラムの構築 I: 新たな価値創造に向けたニーズとシーズのマッチング(PM イノベーション, 日本発信の仕組み P2M コンセプト明確化と発展的展開-イノベーションを促進し, 競争力再生と人材育成に貢献する-), 国際 P2M 学会研究発表大会予稿集, 2009 春季
- [3] 堀江 進之助, 上田 昌平, 佐藤 暢, 野田口 真也, 落合 圭, 山口 祐穂, 坂内 悟, 特許情報と文献情報の統合解析による技術情報の可視化-技術情報から見た高知県及び愛媛県における研究開発及び産学連携の現状とその動向-, 情報プロフェッショナルシンポジウム予稿集, 2009
- [4] 内海 京久, オープン・イノベーションの技術ニーズ・シーズマッチング-産学連携による MEMS 新事業創造事例の分析-, AAOS Transactions J-STAGE 早期公開論文, 2025
- [5] 澤田 芳郎, 大学モデルと産学連携コンフリクト, 産学連携学, 1 巻 1 号 p. 5-8, 2004
- [6] Chuangtao Ma Sriom Chakrabarti Arijit Khan and Bálint Molnár Knowledge Graph-based Retrieval-Augmented Generation for Schema Matching 2025
- [7] Qinggang Zhang Shengyuan Chen Yuanchen Bei Zheng Yuan Huachi Zhou Zijin Hong Junnan Dong Hao Chen Yi Chang Xiao Huang A Survey of Graph Retrieval-Augmented Generation for Customized Large Language Models 2025
- [8] Liu Zhuchenyang AcademicRAG: Knowledge Graph Enhanced Retrieval-Augmented Generation for Academic Resource Discovery 2025
- [9] 奥村拓喜, 佐野睦夫, 生成 AI を用いた企業のニーズと研究室の持つシーズのマッチングシステム, 2025 年度 情報処理学会関西支部 支部大会 講演論文集 (2025.9.28)