

Serendipity Wall: 会話文字起こしのベクター検索と大規模言語モデルによる議論支援システム

今村 翔太^{1,a)} 平城 裕隆^{1,b)} 暦本 純一^{1,2,c)}

概要: グループディスカッションは、新しいアイデアを探求するために重要である。その支援手法のひとつに関連するキーワードや画像の提示がある。しかし、情報の提示に会話の文脈情報が考慮されない傾向があった。そこで本研究では、議論に関連して情報を提示することで、議論を発展させるシステムを提案する。具体例として、HCI 研究者の学術的な議論を取り上げた。議論中、システムは継続的に対話を書き起こし、議論の埋め込みベクトルを生成する。これらのベクトルを既存の研究論文のベクトルと照合し、関連する研究を特定する。そして、LLM が要約しながら、関連する研究を画面に表示する。評価実験において、本システムは議論の話題を広げ、新たな知識の習得を促進する効果があった。本研究は、AI が情報検索、提示、整理による議論支援を行うことで、議論を促進できる可能性を示した。

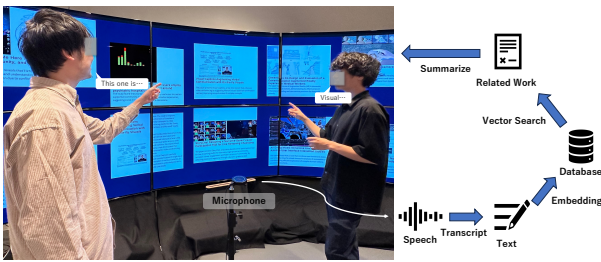


図 1 Serendipity Wall は議論の内容に応じて関連情報を提示することで、議論を促進するシステムである。このシステムは、進行中の議論の参加者の発言をリアルタイムで収集し、書き起こす。書き起こされた文章は時間毎に分割され、埋め込みベクトルに変換される。そして、ベクトル検索によってデータベースから議論に関連する情報を抽出する。抽出された情報は大規模言語モデルによって要約され、ディスプレイに表示される。

1. はじめに

ブレインストーミングとグループディスカッションは、創造的な問題解決とアイデア出しのために不可欠な共同活動であり、製品開発、研究、マーケティング、その他様々な分野で広く使われている [1]。ブレインストーミングセッションは通常、アイデア創出とアイデア共有・増幅の2つのフェーズを含む。他の参加者からのフィードバックや洞察を取り入れることで、集団的な双方向性を育み、グルー

プ全体の創造性を高めることができる。先行研究では、このようなアイデア出しを促進し、活発な議論を促進するための多くのツールが提案されてきている [2], [3]。

自然言語処理技術を用いて、文学やジャーナリズムなど様々な文脈でアイデア創出を支援するデジタルツールも提案されている。大規模言語モデル (以下、LLM) の進歩に伴い、インタラクティブなチャット形式が登場し、SF 小説の冒頭のアイデア [4], [5] やプレスリリースのトピック [6], [7] をユーザに提供している。実用的なアイデア生成タスクでは、Salikutluk ら [8] が、人間がフェルミ推定問題を解いたデータで GPT-3 [9] を訓練することで、LLM が人間の能力に近い推論を生成できることを実証した。また、Summers-Stay ら [10] は、Alternate Users Task のアイデア生成において、LLM が人間の能力を上回ることを示した。また、LLM との対話は、人間のブレインストーミングセッションで観察されるような認知的多様性を示した [11]。さらに、LLM との対話はソーシャルメディア上での態度の変化を示し [12] ており、LLM と人間の共同作業が人間の創造性を高める可能性を示唆している。

議論を活性化する手法として電子的なポストイットによるアイデア出しの支援 [13] やリモート環境でのブレインストーミング [14]、chatbot を用いたファシリテーションによる合意形成の支援 [15], [16], [17], [18] などが提案されてきた。これらのデジタルツールが介入することによってアイデアの創出がより円滑に行われている。中でも、議論中の音声から得られるテキスト情報を用いた議論支援としては、Word2Vec [19] で話題を広げる手法が提案されてい

¹ 東京大学

² ソニーコンピュータサイエンス研究所

a) imamura.shota.ap@alumni.tsukuba.ac.jp

b) hiraki-uts1@g.ecc.u-tokyo.ac.jp

c) rekimoto@acm.org

る [20] ほか、議論によって出てきた単語そのものや web 記事に紐づいた画像を提示することで視覚的に活性化する方法が提案されている [21], [22], [23].

こうした先行研究から LLM はブレインストーミングを拡張して人の認知能力を広めることが期待できる。しかし、先行研究では個別の文書に基づく議論支援が行われておらず、最新の論文のような LLM が知り得ない新たな知見を踏まえて議論した場合の影響が明らかになっていない。また、議論中の会話において既存の研究では単語や Wikipedia データセットを用いた Word2Vec での類義語提示といった単語での支援であり、会話の文脈を踏まえた議論の支援が行われていない。

そこで本研究ではブレインストーミング中の議論を促進するためのシステム、Serendipity Wall を提案する。Serendipity Wall は、進行中の議論の書き起こしをもとに埋め込みベクトルを生成し、関連する情報データベースをベクトル検索することで、議論の内容に近い情報を提示し、議論を促進し、知識を広げる。

HCI 研究の議論による評価実験を行った結果、ユーザーが Serendipity Wall を使って、新しい洞察を生み出したり、関連する画像でディスカッションを促進したり、関連する研究であいまいな関心を明確にしたりすることで、議論を促進できることが確認できた。

2. 関連研究

2.1 大規模言語モデルを活用したアイデア出し支援

先行研究ではデジタルツールを用いて人間の創造性をより高める手法が提案されてきた [2], [3].

GPT3[9] などの自然言語処理の発達によって、大規模言語モデルを用いた創作支援が作曲 [24] やテキスト画像生成 [25], [26], [27] の分野で行われてるようになった。大規模言語モデルは大量のドキュメントを用いて知識を獲得しており、追加の学習をすることなく多くの自然言語処理タスクを解決できる。

Sparks[6] では科学文章の生成を行うためのインスピレーションの提供を提案している。buncho では日本語での小説執筆を支援する。BunCho[4] を使用すると、キーワードからタイトルやあらすじを生成することができ、さらにテーブルトップのロールプレイングゲームとしての対話型ストーリーの制作を行っている。Metaphorian[28] では科学ライターが広い読者層に対して難解な概念をよりアクセスしやすい方法で伝えるための、メタファー作成支援を提案している。

また、近年ではテキストの与え方や例示の仕方によって LLM の出力が改善されることが知られており、プロンプトエンジニアリング [29] と呼ばれている。このような LLM との対話による人間の創造活動は、テキストを用いた様々な活動に広がりつつある。

WORDCRAFT ら [5] は、物語の文章を書く際に人間と LLM が共同で執筆するシステムを提案しており、Wordcraft が提案した案が物語の文章に採用されたことを評価しており、特に文章生成を改善するようなインタフェースも効果があった。

Maurice ら [12] は、言語モデルが特定の意見を支持する場合の人間側の態度の変化について評価を行っている。言語モデルのアシスタントを用いてテキストで議論を行った場合に、被験者のソーシャルメディアでの態度が改善されたことを示した。

ニュースメディアは存在する枠組みや物語を批判しながら、プレスリリースのような文書を基にストーリーのアイデアを探る傾向があり、AngleKindling[7] ではジャーナリストが報道の文章の案を生み出すのに適した方向性を提示した。

2.2 大規模言語モデルを用いたブレインストーミングの支援

創造的なアイデアを生み出すには、異なる視点を考慮する認知的柔軟性が重要と考えられており [30], ブレインストーミングは他者の視点でアイデアを考えられる点で、意思決定の場面において広く利用されている。HCI 研究者によって、議論中に得られる音声データを解析してアイデアの創出を促進するツールが提案されてきた。

IdeaExpander[21] は会話の内容に基づいて動的に画像の刺激を取得し表示することで、視覚的コミュニケーションチャンネルを追加する。IdeaExpander によって、使わない場合に比べてアイデアが平均して 2 つ多く生成された。InspirationWall[22] は、音声データをもとに会話の中から得られたキーワードに基づいて web の情報を検索し、テキストで表示する。Ideawall[23] は、口頭での議論から得られた情報を継続的に抽出し、その情報を Web 検索によって補完する。さらに、情報のビジュアルを視覚的に提示することで議論の活性化を提案している。V8storming[20] は、会話の中からキーワードを抽出し、キーワードに関連する単語を提示することで、プレストにおける話題の幅を広げることを提案している。BrainFax[31] はオンラインホワイトボードアプリの Miro の上でテキストから画像を生成することで、デザインにおけるプレストを活性化している。

ブレインストーミング中の会話を検索するにあたって、これらの研究はキーワードマッチングや Word2Vec[19] といった NLP 手法が使われてきたが、大規模言語モデルの発達によって、LLM を用いたアイデア支援の手法も提案されている。

Idea Machine[32] では入力されたアイデアを LLM によって広げたり、書き直したり、既存の二つのアイデアを組み合わせたりする手法が提案されている。

Summers-Stay ら [8] は、プレストを行う際に GPT-3[9]

に与えるプロンプトによって人間より創造性が上がるかを調査した。発散的思考能力(創造性の重要な側面)のテストである Alternate Uses Task(AUT)[10]のスコアを、LLMとの協調によって向上させられることが明らかになった。

Memmert と Tavanapour[11]は、これまで人間同士で行われてきたプレストで見られていたグループでの認知的刺激の兆候 [33], [34] が、人間と LLM によっても観測できることを示している。

Salikutluk ら [35] は推定問題、つまり”ダッカにはスマートフォンユーザーが何人いますか?”のように不完全な情報から数量を推定する課題を解決するための 議論支援手法を提案した。声に出して考える think-aloud study で観察したところ、人間は推測質問をサブ質問に分解し、意味的に関連した質問に置き換えることが多いことが明らかになった。このデータをもとに GPT-3 を訓練すると GPT3 は人間のように、推定問題のサブクエスチョンやそれに関連する質問を考え出すことができるようになった。

これらの LLM による議論支援は、LLM の中にある情報に依存しているため、論文のような最新のデータに基づいてアイデア出しを行った場合の影響が明らかになっていない。

2.3 ベクター検索を用いた文書情報検索

大規模なドキュメントデータベースから目的のドキュメントを正確かつ高速に見つけることは検索をインタラクティブに行う上で重要である。膨大なデータから検索を行う際には、おおまかに検索して一部を取り出した後に、その中で検索クエリに対して近い順にランキングをつけるといった二段階のプロセスを行う [36]。特に BERT でのランキングが行われてきた [37], [38] が、マッチングアルゴリズムを実行するため計算量が膨大であった。LLM に代表されるように Next sentence prediction や Masked Language M Odelling などの表現学習を用いた言語モデルが提案されることで、ベクトル埋め込みを用いたインデックスによる検索が提案されてきた。文章ベクトルと入力クエリを文章ベクトルに埋め込むと、最近傍探索問題として実行できるほか、また、LLM と最近傍探索の組み合わせは少ないパラメーターのモデルであっても実行でき、追加学習なく性能を発揮できる。このようなベクトル埋め込みを用いた最近傍探索は Faiss[39] や SPTAG[40] が提案されており、要約の生成 [41] にも用いられている。

3. 手法

3.1 システム構成

本研究では議論を支援するために、議論の会話内容に合わせて関連情報を提示する Serendipity Wall を実装した。Serendipity Wall は情報提示を行うための大型ディスプレイ群と、収録のためのマイク、Unity による情報提示アプリ

ケーションのフロントエンドシステムと、情報検索のためのローカルで動かす FastAPI サーバー、関連情報のデータベースとそれらの処理を動かすパソコンから構成される。



図 2 Serendipity Wall は6台の曲面 4K ディスプレイとマイクから構成され、議論の内容に応じた関連情報を提示する。

3.2 文字起こし

マイクで収録された会話の内容は 20 秒ごとに wav ファイルで保存された。保存された音声ファイルは OpenAI 社の提供する Whisper API[42] を用いて文字起こしを行なった。文字起こしを行なったテキストは関連情報を検索するためのほか、評価のための会話テキストとして保存した。

3.3 データベース

本研究では、議論支援の対象として Human Computer Interaction の研究アイデアのディスカッションを例にシステムの実装を行うこととした。提示する関連情報としては、HCI に関する広範囲のテーマをカバーすることから、CHI2023 から 874 件の論文を対象とすることとした。CHI2023 の研究論文に関するデータベースを構築するために、まず、ACM Digital Library より CHI2023[43] のプロシーディングスの PDF をダウンロードした。そして Python を用いて、その PDF のパースを行い、PDF の冒頭にある研究に関連する部分以外については除外した上で、” ABSTRACT ” という文字を含むページで分割を行った。これにより CHI2023 から 874 件の論文の PDF を取得した。

次に、各研究の PDF の最初のページの上半分と、各論文 PDF 画像中で最初に出てくる写真の抽出を行った。なるべく研究を象徴するキービジュアルとなる画像を抽出できるように、抽出した最初のページ上部の画像と論文中の最初に出てくる画像の内、より色彩が豊かな方の画像をキービジュアルとして採用してデータベースに保存した。なお、論文の中で画像を用いていない論文に関しては、そのまま最初のページの上半分をキービジュアルとして取り扱った。



図 3 情報の更新

3.4 埋め込みベクトルとベクター検索

多くの論文文献の中から会話の内容に近い文献を探せるようにするため、埋め込みベクトル [19] を作成した。埋め込みベクトルは自然言語処理で使われる手法の一つで、単語やフレーズを高品質な分散ベクトル表現に変換する。埋め込みベクトルは、各単語や文の意味や文脈を捉えることができる。似たような意味を持つ単語や表現は、埋め込み空間ではより近くにマッピングされる傾向がある。この性質を用いて情報検索においても活用されている。

本研究では、埋め込みベクトルは OpenAI 社の提供する Embedding API[44] を用いて生成した。この API の生成する埋め込みベクトルは 1536 次元となっている。一度に埋め込みベクトルにできる最大トークン数が決まっており、8191 トークンとなっている。これは英文字にして 1000 文字程度である。論文の最初のページにはタイトルやアブストラクト、キーワードやティザー画像のキャプション情報などが入っており、各研究の内容がまとまっていることが期待できる。そこで、本研究の実装では各研究論文の最初の 1 ページを埋め込みベクトルとすることにした。また、最新の会話の方向性を反映しやすくするため、会話の内容の 20 秒毎の文字起こしを一つの会話の埋め込みベクトルとすることにした。

会話の内容に関連度の高い論文を検索するために、FAISS[39] というベクトルの効率的な類似性検索を行えるライブラリを用いてベクトル検索を行った。ベクトル検索はローカルマシン上で稼働させた Python の Web API サーバーである Fast API のコード上で実行された。

3.5 情報の提示

情報の提示は最初は議論のトピック（例えば「AI の活用方法について」など）を入力して出てきたベクター検索結果の上位 10 件の関連研究の表示を行った。その後は 20 秒毎に一件ずつ、会話の埋め込みベクトルに近い関連研究を追加するようにした。追加は左右両端にある計 4 件の表示のうち、最も関連度合いの低いものを押し出すように反対側に新しい研究の情報パネルを挿入して行った (図 3)。

各研究の表示は情報量の異なる二種類のものを用意した (図 4)。一つはキービジュアルとタイトル、大規模言語モデルを用いた短文の要約を表示した。短文要約は論文の最初のページをコンテキストに与えた上で、要点や主張を 30

文字以内に要約するようプロンプトで指示をした。大規模言語モデルには OpenAI 社の ChatGPT API (GPT-4) [45] を用いた。短文の要約だけでは情報が不足するため、もう一つの表示形態として、キービジュアルとタイトル、大規模言語モデルを用いた課題、手法、結果の要約の表示も用意した。

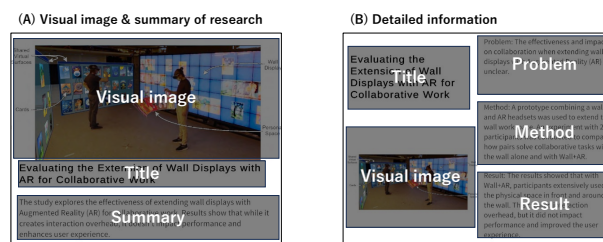


図 4 (A) タイトルと画像を中心にした提示形態。(B) より詳細に説明した提示形態。

詳細の情報の提示はすでに表示されている研究がベクター検索にて再度、最上位に来た場合に行われた。ただし、特定の研究に偏らないよう、3 度目以降の場合には次点以降の研究を表示するようにした。

3.6 装置について

複数人が同時に対面でディスカッションをしながら提示される様々な情報を見ることができるよう、LG 社製 65EG9600 の湾曲 4K 有機 EL ディスプレイ 6 枚を一体的なディスプレイとして組み合わせた大型ディスプレイ空間を実装した。ただし、計算機の負荷低減のために各ディスプレイは HD 画質での出力としたため、全体の解像度は 5760*2160 である。

また、複数人の話者の声を拾えるように、Seed Studio 社製の ReSpeaker USB Mic Array マイクをスタンドの上に設置して使用した。

4. 評価

4.1 評価実験

提案手法による議論支援の効果を検証するために提案手法を用いて HCI 研究に関するディスカッションを行なってもらう実験 (図 5) を実施した。HCI に関連する研究室から 6 名の HCI の研究者及び学生を集めた。同一の研究室から各 2 名ずつの研究者のペアを計 3 ペア作った。いずれの実験参加者も日本語のネイティブ話者であった。この

ため関連情報の要約の提示や議論については日本語で行われた。各ペアは15分間のHCIの研究に関するディスカッションを2回行ってもらった。各ディスカッションにおいては参加者の研究内容や興味関心に応じたトピックを1つずつ設定した。例えば、「モビリティを活用した新しい研究アイデア」といったトピックがあった。一方のディスカッションでは提案手法を用いてトピックに関連する研究を表示した中でディスカッションを行なってもらい、もう一方のディスカッションではトピックのみを決めて自由にディスカッションを行なっていただいた。実験参加者らには、トピックに関連して、より多くのアイデアや広いアイデアについて議論してもらうようにした。特に特定の結論を出したり、批判的な議論を行う必要はないことを説明した上でディスカッションを行なっていただいた。



図 5 2名の実験参加者が Serendipity Wall の前で議論している様子

4.2 評価指標

主観的な評価と会話の埋め込みベクトル、会話のボキャブラリーの多様性の3つの指標をもとに本システムを評価した。

4.2.1 主観的な評価

ディスカッションを終了した後にアンケートに回答していただき、以下の項目について回答を得た。

- Q1. 議論に関連する研究が出てきたか？
- Q2. 関連研究の情報提示は議論の促進に役立ったか？
- Q3. 情報提示により研究アイデアを具体化するためのヒントを得ることはできたか？
- Q4. 情報提示により新しい関連研究に触れることができたか？
- Q5. 情報提示の関連研究の情報量が適切だったか？
- Q6. 情報提示で一度に表示される研究の数が適切だったか？
- Q7. 情報提示ありの場合と情報提示なしの場合でどちらがより議論が盛り上がったか？
- Q8. 関連研究の情報提示が役立った、もしくは役立たなかった理由

Q9. 表示形態や欲しい情報についての意見

4.2.2 会話ベクトルによる評価

会話ベクトルを用いて、議論の広がりについても評価した。まず、評価議論中の発話の文字起こしを20秒毎に埋め込みベクトルにした。次に、一つのディスカッション中の各埋め込みベクトルとのコサイン類似度を計算し、その平均と標準偏差を各手法毎に導出した。このcos類似度の平均と標準偏差をもとに議論の広がり进行评估した。潜在空間の中での変化が分かりやすくなるよう、CHIの論文の埋め込みベクトルの平均をHCIの研究の部分空間であると見做して、会話ベクトルの原点をCHIの論文の埋め込みベクトルの平均へと動かした上でコサイン類似度とその分散の計算を行なった。

4.3 会話のボキャブラリーの多様性の評価

議論に置いて用いられている語彙のボキャブラリーを評価することで議論の特性を評価した。語彙の多様性はMATTR[46]とMTLD[47]の2つの指標を用いて評価した。MATTRのウィンドウサイズは100、MTLDの閾値は0.72に設定した。議論は日本語で行われているため、まず文字起こしを日本語の形態素解析ツールであるMecab[48]で形態素解析を行った。その上で、名詞、動詞、形容詞のみを抽出し、それらをトークンとしてMATTRとMTLDで評価した。

4.4 主観的評価の結果

アンケートのQ1からQ7の回答結果は図6に示した通りであった。

提示されている研究の議論への関連度合いについては6人中6人が関連しているもしくはある程度関連していると回答した(Q1)。また、5名が関連情報の提示が議論の促進に役立ったと回答した(Q2)。関連研究の提示は「新しい話題を話し始めるのに役に立った(P1)」といったように、話題のきっかけとして活用されていた。また、「表示された研究について話しているとしっかりその詳細が出たので、役に立った(P5)」といった感想もあった。また、関連した研究の写真を活用して議論を促進できたというフィードバックもあった。「自分がそれらの研究に関して知識がある場合、話している内容と直接関係がなくても、類似した別の研究の画像などを使って、他者に説明することができた(P4)。「絵が出てきたのが役立った。自分の専門外の分野について、情報提示がなければジェスチャーだけになるところを、絵があることでさらに理解しやすくなった。(P3)」

6人中6人が新しい関連研究に触れることができたと回答した(Q4)。「類似の研究が表示されることで、それまで知らなかったその論文を読んでみようと思った。」「同じトピックの用語をもとに論文がどんどん連想されてよかつ

中立

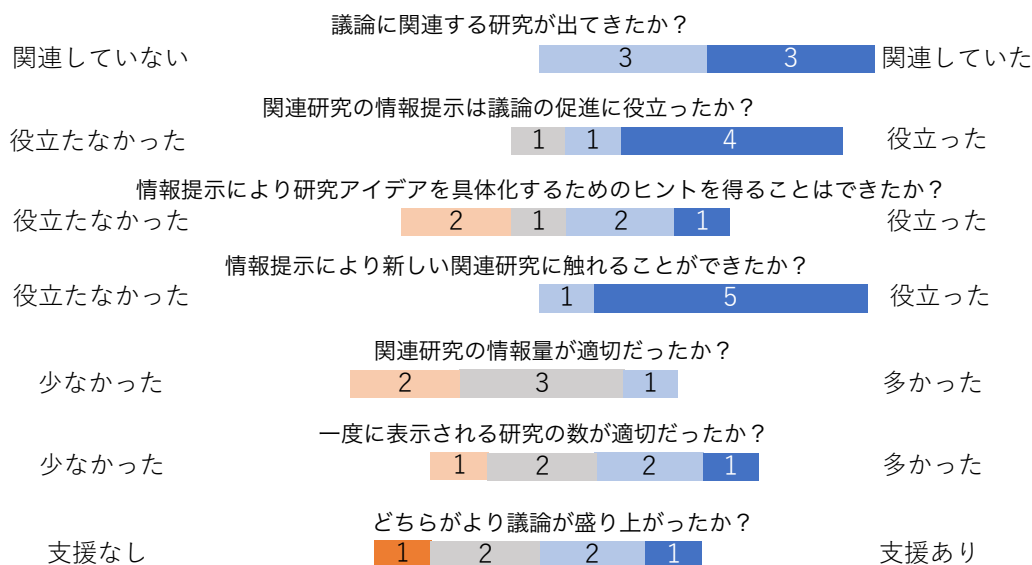


図 6 主観的評価の結果

た。実際に使えそうな関連研究を見つけられた。」といった指摘があった。

一方で、提示情報をもとに、持っているアイデアを具体化していくという点に関しては6人中3人が役立つ、ある程度役立つと回答したが、2人はあまり役立たないと回答した(Q3)。具体化に役立ったと回答した人からは、「補助する存在の特性のような、抽象的な話題の時に、それに関する論文が出てきたことで、それをフックに議論することができ、議論がしやすかった。(P3)」といった理由が上がった。一方で、具体化にはあまり役立たなかった理由としては「議論を深める前に画面が変わるので広く浅く話しがちであった。(P1)」、「知らない研究を知ることができ、より内容を知りたいものがあったが、表示されている内容が限定的で理解を深めることができなかった。(P2)」、「新しい研究がポップアップされるとすぐその話題に行くので、話を深めるという点においては不向きかもしれない。話を発散させるだけでなく、話を深める方向に研究を表示できるともっと良くなりそう。(P5)」といった指摘があった。

総じて関連研究の情報提示のある場合とない場合で、どちらが議論が活発になったかについては、1名が情報提示あり、2名がどちらかといえば情報提示あり、中立が2名、情報提示がないほうが議論が活発になったが1名であった(Q7)。情報提示がないほうがよかったと回答した人はその理由について、「新しい話題を始めるのには役だったが、議論を深める前に画面が切り替わってしまい広く浅く話しがちであった(P1)」ことを理由に挙げた。

提示されている情報量については各研究毎の情報量としてはやや多いが1名、ちょうどいいが3名、やや少ないが2名であった(Q5)。一度に表示される関連研究の数は多い1名、やや多い2名、ちょうど良い1名、やや少ないが1名であった(Q6)。「もう少し短文でまとまっていると見やすい(P1)」、「表示する研究の数を減らして、もう少しテキスト情報を表示きて欲しかったように思いました。もう少し文章量が多い方が個人的には嬉しかった。(P2)」と人によって欲しい情報量には違いがあった。

「自動で表示されるだけでなく、触れると情報をもっと開示できる、スワイプして見飽きた研究を画面外に出すなどインタラクションがあると良いと思った(P2)」といった指摘もあった。

また、関連研究の提示だけでなく、特定の専門単語やコンセプトなどについての解説を加えることへの希望も複数あった。「専門的な装置や指標の話をしている際に、写真、測定する項目など、相手のイメージを補助するものや細かい事項を直接的に提示してくれると、説明が楽になると思った。(P3)」「論文に使われている言葉がそもそもわからなくて、深掘りできない場面があった(P6)」

配置や表示形態に関してのフィードバックもあった。「研究の詳細は、その研究の近くに置いて、ちょっと大きく表示してほしい(P5)」「動画コンテンツなどバリエーションがあっても良い(P4)」

4.5 会話ベクトルによる評価結果

15分間の議論を20秒毎に埋め込みベクトルに変換し、HCIの研究の部分空間に変換したベクトルのcos類似度の平均と標準偏差を導出した(図6)。Serendipity Wallを利用したときは、コサイン類似度の平均は0.495、標準偏差は0.109であった。一方、Serendipity Wallを用いなかった場合には、コサイン類似度の平均は0.522、標準偏差は0.096であった。会話ベクトルが空間中でより広がっていることから、Serendipity Wallによる関連研究の情報提示がある場合には、より多くの研究についての会話が行われていることが窺われる。対して、Serendipity Wallを用いない場合には、会話ベクトルの広がりや狭くなっており、より狭い話題を集中して話している傾向が窺われる。

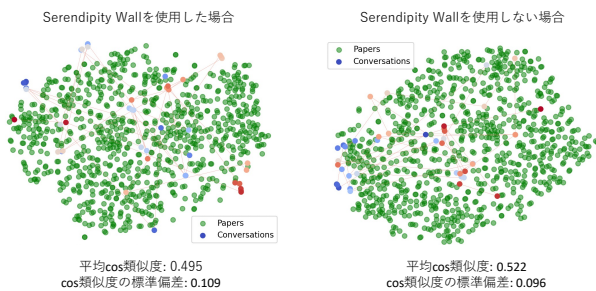


図7 Serendipity Wallを用いた場合と用いなかった場合の各論文と会話のベクトルを次元圧縮して示した。緑色の点はCHI2023の各研究の次元圧縮結果である。それ以外の点は20秒毎の会話のベクトルの次元圧縮結果であり、時系列の変化に合わせて青色から赤色に着色している。

4.6 ボキャブラリーによる評価結果

15分間の議論全体を用いてMATTRとMTLDを導出した。その結果、Serendipity Wallを通じた関連研究の情報提示を行っていた場合には、MATTRは0.685、MTLDは54.8になった。一方で、Serendipity Wallを用いなかった場合には、MATTRは0.715、MTLDは82.6になった。Serendipity Wallを用いない場合には、いずれの実験参加者のペアの場合にも、MATTR、MTLDはより大きな値を示しており、Serendipity Wallを用いないほうが会話に用いられる語彙の多様性があることが窺われる。

5. 議論

5.1 議論支援による議論の広がり

実験参加者のフィードバックやベクターの広がりからは、Serendipity Wallを通じて議論内容に関連した研究の要約を提示することで、話のきっかけが生まれたり、関連した研究やアイデアを手に入れることができ、HCIの研究に関する議論が拡大していることがわかった。また、曖昧なコンセプトから関連した研究をもとに議論の具体化をしたり、写真を活用したコミュニケーションを図ったりと

いった議論の促進効果も見られた。

一方で、現在のシステムは、一度は議論の内容に合わせて研究の詳細について提示するものの、その後は次々と他の関連した研究を提示していく実装になっており、話が発散してしまったり、議論を深ぼることが難しいといった欠点があることが明らかになった。様々な研究に触れたり、アイデアを入手するのか、アイデアを深めていきたいのかといった議論の目的によっても必要な支援が変わってくる。提示処理や形態にバリエーションを持たせるといったシステムの改善が必要であると考えられる。また、実験の議論テーマの設定を工夫して議論の発散や収束のフェーズを再現することで、各提示処理や形態を評価していくことも考えられる。

特に特定研究や特定の用語についての詳細が知りたい場合に関しては、単に議論の内容に関連した研究を提示するだけではなく、会話の内容から他の研究を見たいのか、特定の研究や概念について掘り下げていきたいのかを判定する処理を追加することが考えられる。これにより他の関連研究を提示するのか、特定の研究のさらに知りたい部分の詳細については、特定の論文内のベクター検索や大規模言語モデルの要約を通じて提示してあげるのかといった処理の切り分けを行うことができるようになる。あるいは深掘りした情報と発散させる情報の両方を表示し、議論の変遷を元にして提示していく情報のバランスを調整することも考えられる。また、これまでの議論を大規模言語モデルを利用して整理したり、アイデア出しをするといった支援のあり方も考えられる。

さらに、評価実験では特定の用語や装置の説明が難しかったり、用語が分からないために議論が深まらなかったといった例があった。関連研究以外にも関連情報を提示することで、議論の促進がより図れたといったフィードバックがあった。関連研究データベース以外にも、必要に応じて広くインターネットの情報を検索したり、大規模言語モデルを用いて説明をこなったりするといった機能を実装することができれば、さらなる議論の促進につながる可能性がある。

また、本研究では、関連研究の提示により議論支援を行なった場合の方が語彙の多様性が低下するといった結果となった。議論は広く浅くなった傾向が定性的なフィードバックや埋め込みベクトルからは示されているが、語彙の多様さといった観点ではトピックのみを決めて自由にディスカッションを行う方が広くなるということになる。これは関連情報の提示により、触れられる研究の内容は広がっているものの、使用される単語は偏っているということを示唆している。

5.2 情報の配置と更新について

本研究の実装では簡単のために、左右から関連度の低い

情報を押し出す形で新しい情報を差し込むといった実装になっている。そのため、関連研究の詳細を提示が、元の関連研究の詳細の位置と離れてしまった場合もあった。情報の関連度に応じて隣接関係がわかる配置にしたり、追加の関連情報は隣接して表示することで、分かりやすさを向上させることができると考えられる。

また、情報の更新頻度も 20 秒に一回一つずつと一定の速度で行われていた。情報の更新頻度や、埋め込みベクトルを作成する会話の時間単位が 20 秒で良いのか、一定の速度で良いのかについても調整や評価が必要であると考えられる。

また、発話内容に関連した情報提示を通じてブレインストーミングを促進する研究にはキーワードや写真のみを提示する先行研究もある。Serendipity Wall では情報の要約など情報量を多く提示するといったシステムになっている。情報量の寡多や提示情報の種類によって議論にどのような変化を及ぼしうることについて調査する必要がある。

5.3 入力の多様化

本研究においては会話の音声の文字起こしのみを入力に使用している。しかし、音声の文字起こしにはいくつか限界がある。まず、発話していない場合に関心を特定することが困難である。加えて文字起こしの精度が特に専門用語や固有名詞などの場合には完璧ではない場合も多い。さらに会話中にある「あれ」や「それ」といった指示語を特定することが音声文字起こしだけでは困難である。そこで会話に参加している視線をトラッキングして、注視対象を元に関心情報の入力を補うというアプローチが考えられる。例えば、視線の注視時間や領域を元に、注視対象の埋め込みベクトルに重み付けを行い、会話の埋め込みベクトルと合わせて、関心情報の埋め込みベクトルを構成するといった実装がありえる。

音声情報と視線情報を組み合わせたマルチモーダルな入力により、関連情報の推薦精度をあげたり、発言をしていない参加者の意図も汲み取ったりすることでより良い議論支援につながる可能性がある。また、一人で利用している場合のアイデア出しや情報探索にも活用しやすいアプリケーションになることも考えられる。

加えて、どこを見ているのかといった情報を入手することができれば、見られている場所の周辺に情報を提示していくといった実装も可能となる。視線情報を活用して、情報の可視化やレイアウトの工夫にも活用していける可能性がある。

5.4 研究以外の用途への適用拡大と装置の多様化

本研究における評価は研究ディスカッションのみをテーマで行われた。しかし、ブレインストーミングやグループディスカッションによる課題解決やアイデア出しというの

は多様な分野で重要となる共同作業形態であり、他の分野での応用を模索していくことが重要である。研究以外の用途に活用する場合に、どういった情報を提示するのかや、ディスカッションの影響に変化がありうるのかといった部分に関して評価を行うことが必要であると考えられる。

また、Serendipity Wall の実装と評価は大型ディスプレイ空間を用いて行われた。より小型のディスプレイを用いた場合にどのような違いが出るのかについて評価することも必要である。また、ビデオ会議の内容に合わせて関連情報を提示するといった用途も考えられる。さらには情報を広く空間に提示するという点でいくと、仮想現実や拡張現実のヘッドマウントディスプレイを通じた共創作業における活用も見込むことができる。こうした多様な形態における情報提示形態の模索や、共創作業への影響について評価していくことで、本研究の成果をより広い範囲で活用していける可能性がある。

6. おわりに

本研究では、議論の内容に応じて関連情報を提示することで議論の促進を図るためのシステム、Serendipity Wall を提案した。ケーススタディでは Serendipity Wall を用いて HCI の研究ディスカッションを行った結果、新たな知見を得たり、議論の円滑化を図ったり、曖昧な関心を具体化したりすることで議論が活発になったり、話が広がる効果があることを確認した。一方で、情報提示の仕方によっては議論の進展を深めることを阻害してしまうという副作用があることも明らかになった。今後、情報の内容や出し方を工夫したり、適応範囲を拡大していくことで、会話の内容に合わせて情報を提示し、議論や情報の取得を促進するアプリケーションを実現していけると考える。

謝辞

本研究は JST ムーンショット型研究開発事業グラント番号 JPMJMS2012, JST CREST グラント番号 JP-MJCR17A3, 国立研究開発法人情報通信研究機構の委託研究 02901 の支援を受けたものです。

参考文献

- [1] Alex Osborn. *Applied Imagination: Principles and Procedures of Creative Problem Solving*. Scribner, New York, 1953.
- [2] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. Mapping the landscape of creativity support tools in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, number Paper 389 in CHI '19, pages 1–18, New York, NY, USA, May 2019. Association for Computing Machinery.
- [3] Ben Shneiderman. Creativity support tools: accelerating discovery and innovation. *Commun. ACM*, 50(12):20–32, December 2007.

- [4] Hiroyuki Ozone, Jun-Li Lu, and Yoichi Ochiai. BunCho: AI supported story Co-Creation via unsupervised multi-task learning to increase writers' creativity in Japanese. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, number Article 19 in CHI EA '21, pages 1–10, New York, NY, USA, May 2021. Association for Computing Machinery.
- [5] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. Wordcraft: Story writing with large language models. In *27th International Conference on Intelligent User Interfaces, IUI '22*, pages 841–852, New York, NY, USA, March 2022. Association for Computing Machinery.
- [6] Katy Ilonka Gero, Vivian Liu, and Lydia Chilton. Sparks: Inspiration for science writing using language models. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference, DIS '22*, pages 1002–1019, New York, NY, USA, June 2022. Association for Computing Machinery.
- [7] Savvas Petridis, Nicholas Diakopoulos, Kevin Crowston, Mark Hansen, Keren Henderson, Stan Jastrzebski, Jeffrey V Nickerson, and Lydia B Chilton. AngleKindling: Supporting journalistic angle ideation with large language models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, number Article 225 in CHI '23, pages 1–16, New York, NY, USA, April 2023. Association for Computing Machinery.
- [8] Douglas Summers-Stay, Clare R. Voss, and Stephanie M. Lukin. Brainstorm, then select: a generative language model improves its creativity score. In *The AAAI-23 Workshop on Creative AI Across Modalities*, 2023.
- [9] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901. Curran Associates, Inc., 2020.
- [10] Christensen P. R. Merrifield P. R. Wilson R. C. Guilford, J. P. Alternate uses (altus). *APA PsycTests*, 1960.
- [11] Lucas Memmert and Navid Tavanapour. TOWARDS HUMAN-AI-COLLABORATION IN BRAINSTORMING: EMPIRICAL INSIGHTS INTO THE PERCEPTION OF WORKING WITH a GENERATIVE AI. page 429, 2023.
- [12] Maurice Jakesch, Advait Bhat, Daniel Buschek, Lior Zalmanson, and Mor Naaman. Co-Writing with opinionated language models affects users' views. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, number Article 111 in CHI '23, pages 1–15, New York, NY, USA, April 2023. Association for Computing Machinery.
- [13] Scott R Klemmer, Mark W Newman, Ryan Farrell, Mark Bilezikjian, and James A Landay. The designers' outpost: a tangible interface for collaborative web site. In *Proceedings of the 14th annual ACM symposium on User interface software and technology, UIST '01*, pages 1–10, New York, NY, USA, November 2001. Association for Computing Machinery.
- [14] Khanh-Duy Le, Paweł W Woźniak, Ali Alavi, Morten Fjeld, and Andreas Kunz. DigiMetaplan: supporting facilitated brainstorming for distributed business teams. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*, number Article 36 in MUM '19, pages 1–12, New York, NY, USA, November 2019. Association for Computing Machinery.
- [15] Joongi Shin, Michael A Hedderich, Andrés Lucero, and Antti Oulasvirta. Chatbots facilitating Consensus-Building in asynchronous Co-Design. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, number Article 78 in UIST '22, pages 1–13, New York, NY, USA, October 2022. Association for Computing Machinery.
- [16] Sangah Park, Yoon Young Lee, Soobin Cho, Minjoon Kim, and Joongseek Lee. “knock knock, here is an answer from next door”: Designing a knowledge sharing chatbot to connect residents: Community chatbot design case study. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing, CSCW '21*, pages 144–148, New York, NY, USA, October 2021. Association for Computing Machinery.
- [17] Donghoon Shin, Soomin Kim, Ruoxi Shang, Joonhwan Lee, and Gary Hsieh. IntroBot: Exploring the use of chatbot-assisted familiarization in online collaborative groups. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, number Article 613 in CHI '23, pages 1–13, New York, NY, USA, April 2023. Association for Computing Machinery.
- [18] Amon Rapp, Arianna Boldi, Lorenzo Curti, Alessandro Perrucci, and Rossana Simeoni. Collaborating with a Text-Based chatbot: An exploration of Real-World collaboration strategies enacted during Human-Chatbot interactions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, number Article 115 in CHI '23, pages 1–17, New York, NY, USA, April 2023. Association for Computing Machinery.
- [19] Tomas Mikolov, Kai Chen, G.s Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *Proceedings of Workshop at ICLR*, 2013, 01 2013.
- [20] Yui Kita and Jun Rekimoto. V8 storming: How far should two ideas be? In *Proceedings of the 9th Augmented Human International Conference*, number Article 14 in AH '18, pages 1–8, New York, NY, USA, February 2018. Association for Computing Machinery.
- [21] Hao-Chuan Wang, Dan Cosley, and Susan R Fussell. Idea expander: supporting group brainstorming with conversationally triggered visual thinking stimuli. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work, CSCW '10*, pages 103–106, New York, NY, USA, February 2010. Association for Computing Machinery.
- [22] Salvatore Andolina, Khalil Klouche, Diogo Cabral, Tuukka Ruotsalo, and Giulio Jacucci. InspirationWall: Supporting idea generation through automatic information exploration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition, C&C '15*, pages 103–106, New York, NY, USA, June 2015. Association for Computing Machinery.
- [23] Yang Shi, Yang Wang, Ye Qi, John Chen, Xiaoyao Xu, and Kwan-Liu Ma. IdeaWall: Improving creative collaboration through combinatorial visual stimuli. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW*

- '17, pages 594–603, New York, NY, USA, February 2017. Association for Computing Machinery.
- [24] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. Novice-ai music co-creation via ai-steering tools for deep generative models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–13, New York, NY, USA, 2020. Association for Computing Machinery.
- [25] Vivian Liu and Lydia B Chilton. Design guidelines for prompt engineering text-to-image generative models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.
- [26] Vivian Liu, Han Qiao, and Lydia Chilton. Opal: Multi-modal image generation for news illustration. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, UIST '22, New York, NY, USA, 2022. Association for Computing Machinery.
- [27] Sitong Wang, Savvas Petridis, Taeahn Kwon, Xiaojuan Ma, and Lydia B Chilton. Popblends: Strategies for conceptual blending with large language models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, New York, NY, USA, 2023. Association for Computing Machinery.
- [28] Jeongyeon Kim, Sangho Suh, Lydia B Chilton, and Haijun Xia. Metaphorian: Leveraging large language models to support extended metaphor creation for science writing. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, DIS '23, pages 115–135, New York, NY, USA, July 2023. Association for Computing Machinery.
- [29] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Comput. Surv.*, 55(9), jan 2023.
- [30] Bernard A. Nijstad, Carsten K. W. De Dreu, Eric F. Rietzschel, and Matthijs Baas. The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *European Review of Social Psychology*, 21(1):34–77, 2010.
- [31] Mathias Peter Verheijden and Mathias Funk. Collaborative diffusion: Boosting designerly Co-Creation with generative AI. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, number Article 73 in CHI EA '23, pages 1–8, New York, NY, USA, April 2023. Association for Computing Machinery.
- [32] Giulia Di Fede, Davide Rocchesso, Steven P Dow, and Salvatore Andolina. The idea machine: LLM-based expansion, rewriting, combination, and suggestion of ideas. In *Proceedings of the 14th Conference on Creativity and Cognition*, C&C '22, pages 623–627, New York, NY, USA, June 2022. Association for Computing Machinery.
- [33] Alain Pinsonneault, Henri Barki, R Brent Gallupe, and Norberto Hoppen. Electronic brainstorming: The illusion of productivity. *Information Systems Research*, 10(2):110–133, June 1999.
- [34] J F Nunamaker, Alan R Dennis, Joseph S Valacich, Douglas Vogel, and Joey F George. Electronic meeting systems. *Commun. ACM*, 34(7):40–61, July 1991.
- [35] Vildan Salikutluk, Dorothea Koert, and Frank Jäkel. Interacting with large language models: A case study on AI-aided brainstorming for guesstimation problems. In *Frontiers in Artificial Intelligence and Applications*, volume 368 of *Frontiers in artificial intelligence and applications*, pages 153 – 167. HHAI 2023 Augmenting Human Intellect, June 2023.
- [36] Jiafeng Guo, Yinqiong Cai, Yixing Fan, Fei Sun, Ruqing Zhang, and Xueqi Cheng. Semantic models for the First-Stage retrieval: A comprehensive review. *ACM Trans. Inf. Syst. Secur.*, 40(4):1–42, March 2022.
- [37] Rodrigo Nogueira and Kyunghyun Cho. Passage re-ranking with BERT. January 2019.
- [38] Jiafeng Guo, Yixing Fan, Qingyao Ai, and W Bruce Croft. A deep relevance matching model for ad-hoc retrieval. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, CIKM '16, pages 55–64, New York, NY, USA, October 2016. Association for Computing Machinery.
- [39] Jeff Johnson, Matthijs Douze, and Hervé Jégou. Billion-Scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3):535–547, July 2021.
- [40] Qi Chen, Haidong Wang, Mingqin Li, Gang Ren, Scarlett Li, Jeffery Zhu, Jason Li, Chuanjie Liu, Lintao Zhang, and Jingdong Wang. *SPTAG: A library for fast approximate nearest neighbor search*, 2018.
- [41] Xiaoyu Zhang, Jianping Li, Po-Wei Chi, Senthil Chandrasegaran, and Kwan-Liu Ma. ConceptEVA: Concept-Based interactive exploration and customization of document summaries. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, number Article 204 in CHI '23, pages 1–16, New York, NY, USA, April 2023. Association for Computing Machinery.
- [42] OpenAI. Introducing chatgpt and whisper apis, 2022.
- [43] *CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2023. Association for Computing Machinery.
- [44] OpenAI. Embeddings - openai api, 2022.
- [45] OpenAI. Gpt models - openai api, 2023.
- [46] Michael A. Covington and Joe D. McFall. Cutting the gordian knot: The moving-average type-token ratio (mattr). *Journal of Quantitative Linguistics*, 17(2):94–100, 2010.
- [47] Philip M McCarthy and Scott Jarvis. Mtdl, vocd-d, and hd-d: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior research methods*, 42(2):381–392, 2010.
- [48] T. KUDO. Mecab : Yet another part-of-speech and morphological analyzer. <http://mecab.sourceforge.net/>, 2005.