

Title	【課題研究報告書】オンライン議論における注目コンテンツ抽出
Author(s)	王, 治中
Citation	
Issue Date	2024-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18905
Rights	
Description	Supervisor: 長谷川 忍, 先端科学技術研究科, 修士(情報科学)

課題研究報告書

オンライン議論における注目コンテンツ抽出

王 治中

主指導教員 長谷川 忍

北陸先端科学技術大学院大学
先端科学技術研究科
(情報科学)

令和6年3月

概要

合意とは、二人以上の当事者が共通の意見や解決策、決定に至るプロセス、またはその結果である。社会的、経済的、政治的な文脈において、合意形成はグループや組織、コミュニティが前進するための基本的なメカニズムである。合意に至ることは、対立や異なる視点を持つ個人間で共通の土台を見出し、共有の目標に向かって協力することを可能にする。

したがって、効果的な合意形成は、単に決定を下すこと以上の価値を持つ。それは、関係者間の関係を強化し、将来的な協力の基盤を築き、より良い結果をもたらすプロセスそのものである。

会議における合意形成は、参加者が共通の目標や解決策に向けて協力し、共有の理解や決定に至るプロセスである。このプロセスは、異なる背景や視点を持つ人々が一堂に会し、情報を共有し、意見を交換し、議論を深める会議の環境において特に重要である。会議における合意形成は、組織やコミュニティの目標達成、問題解決、方針決定など、さまざまな目的で利用される。そのプロセスには、通常、以下のステップが含まれる。まず、問題の明確化や目標の設定が行われ、参加者がその範囲内で意見を共有する。次に、提案された解決策やアイデアに対して議論が行われ、参加者はそれぞれの利点や欠点、実行可能性などについて意見を交換する。議論を通じて、参加者は相互に理解を深め、異なる視点を考慮した上で、共通の解決策や決定に近づいていく。最終的には、全員が支持できる合意点や妥協点が見つかり、共有の決定が形成される。

それを実現するために、効果的なコミュニケーション、オープンな議論、相互尊重、柔軟性が必要である。参加者が自らの意見を自由に表明し、他者の意見に耳を傾け、必要に応じて自らの立場を調整できる環境が、合意形成を促進する。また、ファシリテーターの役割も重要であり、議論を適切に導き、対立を管理し、参加者が共通の目標に焦点を当てられるよう支援する。しかし、合意形成は常に簡単なわけではなく、時には強い意見の相違や対立が障壁となることがある。このような状況を解決するために、ファシリテーターが必要である。

ファシリテーターは会議における合意形成プロセスにおいて中核となる役割を担っている。彼らの主な任務は、議論を構造化し、全ての参加者が意見を表明し、聞かれる機会を持つよう促進することである。ファシリテーターは、対立が生じた際には中立的な立場を保ち、建設的な解決策を模索することで対立を管理する。さらに、参加者間のコミュニケーションを促し、相互理解を深める役割も果たす。合意形成を促進するために、ファシリテーターは共通の理解や妥協点に向けて議論を導き、共通の目標や解決策に焦点を当てる。また、会議の内容や決定事項、未解決の問題を記録し要約することで、参加者が議論の結果を理解し、次のステップに進むのを支援する。

新型コロナウイルスのパンデミックは会議形式に顕著な影響を及ぼし、多くの組織やグループが対面会議からオンライン会議への移行を余儀なくされた。この急速な変化は、会議の運営、参加者間のコミュニケーション、および合意形成のプロセスにおいて、複数の重要な影響をもたらしている。オンライン会議の導入は、地理的な制約をなくし、より多様な参加者が容易に会議に参加できるようになるなど、参加とアクセスのしやすさが向上した。オンライン会議を効果的に運営し、合意形成を促進するためには、参加者が技術的な問題を克服し、オンライン環境で積極的に関与できるようにする支援とともに、セキュリティとプライバシーを保護するための措置を講じることが必要である。オンライン会議の利点を最大限に活用し、新たな課題に対処することで、パンデミック下でも質の高い議論と効果的な合意形成が実現される。

オンライン会議での合意形成を支援するために、AI エージェントの開発が進んでいる。これらのエージェントは、議論の構造化、情報の整理と要約、参加者の参加促進、対立と問題の管理、合意形成の促進、および技術的サポートなど、多岐にわたる機能を提供する。エージェントの目的は、オンライン会議特有の課題に対処し、会議の効率と生産性を高めることにある。議論を焦点化し、参加者が積極的に関与することを奨励し、建設的な解決策へと導くことで、エージェントは会議の成果を最大化するための重要な役割を果たす。これらのエージェントの開発と導入により、オンライン会議の課題を克服し、より生産的で満足度の高い会議の実現が期待される。

ファシリテーターは会議の進行を円滑にし、効果的な合意形成を促進する重要な役割を担っているが、その過程で意図せずに主観的な影響を与える可能性がある。このような主観性は、会議の結果に偏りを生じさせ、すべての参加者の意見が平等に扱われない原因となることがある。それを解決するために、本研究では、会議中の客観的な注目内容を抽出するためのエージェントを提案した。このエージェントは、会議の進行においてファシリテーターの主観的影響を軽減し、参加者からの意見や提案を公平かつ客観的に評価することを目的として、会議の効率性と公正性が向上し、より有意義な合意形成が促進されることが期待できる。その仕組みとして、オンライン会議中に発生するテキストデータを収集し、自然言語処理技術を用いてトピック抽出や注目度判定を行った。続いて、エージェントを評価するために、オンライン会議用のプラットフォームを構築し、その上で会議の実験を行い、エージェントに対する評価を行った。

実験では、本学の複数の領域に所属している学生を被験者として集め、持続可能な開発目標の問題をテーマとして、同じ日の午前と午後分けて1回の実験を30分として計6回実施した。そして、実験終了後にアンケート調査を行った。アンケートでは、参加者が10分ごとに注目した内容を選択する形式で行われ、多くの参加者が提供された選択肢の中から自分が注目した内容を見つけることができた。しかし、一部の参加者については「その他」を選択し、アンケートの選択肢が全ての注目内容をカバーしていないことが示唆された。今後の研究で注目内容の精

度を向上させるために、参加者が入力した内容を分析する必要があることが示された。

本研究で開発された AI エージェントは、オンライン会議における合意形成を支援する有効なツールであることが示された。今後の研究では、さらに多様な会議環境や議論テーマにおけるエージェントの適用性と効果を検証し、オンライン会議の合意形成をより効果的に支援する方法を探求することが期待される。

Abstract

Consensus is the process or outcome where two or more parties reach a common opinion, solution, or decision. In social, economic, and political contexts, consensus-building is fundamental for groups, organizations, and communities to move forward. Achieving consensus allows individuals with conflicts or differing perspectives to find common ground and cooperate toward shared goals.

Effective consensus building holds value beyond merely making decisions. It strengthens relationships among stakeholders, lays the foundation for future cooperation, and leads to better outcomes.

In meetings, consensus building involves participants working together towards common goals or solutions and reaching a shared understanding or decision. This process is particularly important in meeting environments where people from diverse backgrounds and perspectives come together to share information, exchange opinions, and deepen discussions. Consensus building in meetings serves various purposes, such as achieving organizational or community objectives, solving problems, and making policy decisions. Typically, the process includes clarifying issues or setting goals, discussing proposed solutions or ideas, and exchanging views on their merits, drawbacks, and feasibility. Through discussion, participants enhance mutual understanding and gradually converge on a common solution or decision, eventually finding an agreement or compromise point everyone can support.

Effective communication, open discussions, mutual respect, and flexibility are essential for realizing this. An environment where participants can freely express their opinions, listen to others, and adjust their positions as necessary promotes consensus building. The role of facilitators is also crucial in guiding discussions, managing conflicts, and helping participants focus on common goals. However, consensus building is not always easy, and strong differences of opinion or conflicts can sometimes be barriers. Facilitators are needed to resolve such situations.

Facilitators play a central role in the consensus-building process in meetings. Their main task is to structure discussions and ensure that all participants have the opportunity to express and be heard. They maintain a neutral stance in conflicts and seek constructive solutions to manage them. They also foster communication among participants and deepen mutual understanding. Facilitators guide discussions toward common understanding or compromise points to promote consensus building and focus on shared goals or solutions. They also record and summarize meeting content, decisions, and unresolved issues, helping participants understand the discussion outcomes and move forward.

To effectively manage online meetings and promote consensus building, it is necessary to support participants in overcoming technical issues, actively engaging in the online environment, and taking measures to protect security and privacy. High-quality discussions and effective consensus-building can be achieved even under pandemic conditions by leveraging the advantages of online meetings and addressing new challenges.

The development of AI agents to support consensus-building in online meetings is advancing. These agents offer various functions, including structuring discussions, organizing and summarizing information, encouraging participant engagement, managing conflicts and issues, promoting consensus building, and providing technical support. These agents aim to address the unique challenges of online meetings and enhance the efficiency and productivity of meetings. Agents are crucial in maximizing meeting outcomes by focusing discussions, encouraging active participation, and guiding toward constructive solutions. The development and implementation of these agents are expected to overcome the challenges of online meetings and realize more productive and satisfying meetings.

This study proposed an agent for extracting objective content during meetings to address the potential influence of facilitators on meeting outcomes. This agent aims to reduce the subjective impact of facilitators on meeting proceedings and evaluate opinions and proposals from participants in a fair and objective manner, thereby improving the efficiency and fairness of meetings and promoting more meaningful consensus building. The agent was realized by collecting text data generated during online meetings and applying natural language processing technologies for topic extraction and attention assessment. Subsequently, a platform for online meetings was constructed to conduct meeting experiments and evaluate the agent.

In the experiments, students from various fields within the university were gathered to discuss issues related to the Sustainable Development Goals, conducting six experiments divided into morning and afternoon sessions, each lasting 30 minutes. After the experiments, a survey was conducted. In the survey, participants selected the content they focused on every 10 minutes in a format that allowed them to choose from the provided options. Many participants were able to find the content they focused on among the provided options. However, some participants chose "Other," suggesting that the survey options did not cover all the content of interest. Future research needs to analyze the content entered by participants to improve the accuracy of the focused content.

The AI agent developed in this study has been shown to be an effective tool for supporting consensus-building in online meetings. Future research is expected to explore the applicability and effectiveness of the agent in various meeting envi-

ronments and discussion themes and to seek ways to support consensus building more effectively in online meetings.

目次

目次	i
図目次	iv
表目次	v
アルゴリズム目次	vi
第1章 序論	1
1.1 研究背景	1
1.2 研究目的	2
1.3 論文の構成	2
第2章 関連研究	4
2.1 会議と合意形成	4
2.2 オンライン会議とファシリテーター	5
2.3 従来の会議促進エージェント	7
2.3.1 Kinoshita らの研究	7
2.3.2 Yoshimura らの研究	7
2.3.3 Ito らの研究	8
2.3.4 Okuhara らの研究	8
2.3.5 Yang らの研究	9
2.3.6 Fujita らの研究	9
2.4 本研究の位置付け	10
第3章 ファシリテーター支援エージェント	11
3.1 オンライン会議の構造	11
3.2 エージェントの役割	12
3.3 エージェント処理の流れ	12
3.4 トピック抽出	13
3.4.1 LDA モデルの概要	14
3.4.2 LDA モデルのプロセス	15
3.4.3 LDA モデルの応用	15

3.5	ポスト注目度算出	16
3.5.1	PageRank とは	16
3.5.2	PageRank アルゴリズムの説明	17
3.6	注目内容のトピック判別	18
第4章	実験と考察	20
4.1	実験概要	20
4.1.1	実験目的	20
4.1.2	実験対象	20
4.2	実験方法	20
4.2.1	実験環境	20
4.2.2	実験の流れ	21
4.2.3	実験アンケート	21
4.3	実験システム	21
4.3.1	オンライン会議実験用プラットフォーム	21
4.3.2	トピック抽出機能	22
4.3.3	ポスト注目度算出機能	23
4.3.4	注目ポストのトピック判別と LDA モデルに基づくサマリー生成	24
4.4	実験結果	25
4.4.1	基礎データ	26
4.4.2	トピック抽出機能評価	26
4.4.3	ポスト注目度算出機能評価	29
4.4.4	アンケート評価	34
4.5	考察	37
4.5.1	本研究の知見	37
4.5.2	残された課題	37
第5章	まとめと今後の課題	38
5.1	まとめ	38
5.2	今後の課題	38
5.2.1	大人数の会議実験	38
5.2.2	エージェント処理速度の向上	39
5.2.3	ポストに対する重みつけ	39
5.2.4	ファシリテータでの評価	39
	研究業績	40
	謝辞	41

目 次

1.1	Enter Caption	1
1.2	システムフローチャット	3
3.1	オンライン会議の構造	12
3.2	システムフローチャット	13
3.3	LDA 説明図 [21]	15
3.4	LDA 処理文の例	16
3.5	例文の処理結果	16
3.6	オンライン会議におけるポストの構造例	18
4.1	アンケート結果	35
4.2	全体のアンケート結果	36
4.3	最も注目した内容のアンケート結果	36

表 目 次

4.1	実験用プラットフォーム	22
4.2	実行必要なライブラリのバージョン	22
4.3	実験データ	26
4.4	実験1の最適なLDAモデルパラメータ	27
4.5	実験2の最適なLDAモデルパラメータ	27
4.6	実験3の最適なLDAモデルパラメータ	27
4.7	実験4の最適なLDAモデルパラメータ	28
4.8	実験5の最適なLDAモデルパラメータ	28
4.9	実験6の最適なLDAモデルパラメータ	28
4.10	実験1先頭10分間の注目トピック抽出結果とサマリー	29
4.11	実験1中間10分間の注目トピック抽出結果とサマリー	29
4.12	実験1最後10分間の注目トピック抽出結果とサマリー	30
4.13	実験2先頭10分間の注目トピック抽出結果とサマリー	30
4.14	実験2中間10分間の注目トピック抽出結果とサマリー	30
4.15	実験2最後10分間の注目トピック抽出結果とサマリー	31
4.16	実験3先頭10分間の注目トピック抽出結果とサマリー	31
4.17	実験3中間10分間の注目トピック抽出結果とサマリー	31
4.18	実験3最後10分間の注目トピック抽出結果とサマリー	31
4.19	実験4先頭10分間の注目トピック抽出結果とサマリー	32
4.20	実験4中間10分間の注目トピック抽出結果とサマリー	32
4.21	実験4最後10分間の注目トピック抽出結果とサマリー	32
4.22	実験5先頭10分間の注目トピック抽出結果とサマリー	32
4.23	実験5中間10分間の注目トピック抽出結果とサマリー	33
4.24	実験5最後10分間の注目トピック抽出結果とサマリー	33
4.25	実験6先頭10分間の注目トピック抽出結果とサマリー	33
4.26	実験6中間10分間の注目トピック抽出結果とサマリー	34
4.27	実験6最後10分間の注目トピック抽出結果とサマリー	34

アルゴリズム目次

4.3.1 Discourse フォーラムトピック分析	23
4.3.2 PageRank 計算および上位 3 位の出力	24
4.3.3 LDA モデルと PageRank に基づく会議内容のサマリー生成	25

第1章 序論

1.1 研究背景

会議における合意形成は、参加者間での共通の理解や意見の形成を通じて、問題解決や意思決定を行う重要なプロセスである [1]。このプロセスは、多様な視点や情報を統合し、効果的な決定を導くために不可欠である。特に、ビジネスや学術の分野において、合意形成はプロジェクトの成功や研究の進展に直接的な影響を与える。ファシリテーターは、会議の進行を円滑にし、効果的な合意形成を促進するための重要な役割を担う。彼らは、参加者の意見を引き出し、議論を整理し、合意に向けて導くことで、会議の成果を最大化する。しかし、Griffith らの研究によれば、ファシリテーターの主観が合意形成に影響を与える可能性があり、その経験や公平性が重要な要素となる。 [2]

Ishizuka らの研究では、AI エージェントを利用してオンライン会議における意見の偏りを解決する方法を提案している。 [3] このアプローチでは、AI エージェントが少数派の意見を強調し、よりバランスの取れた議論を促進することを目指している。しかし、この方法の有効性に関する評価はまだ十分ではなく、AI エージェントの介入が会議のダイナミクスにどのような影響を与えるかは明らかにされていない。

一方で新型コロナウイルスの流行により、対面会議からオンラインやハイブリッド会議への移行が加速した。オンライン会議は、地理的な制約を超えた参加やコスト効率の向上をもたらすが、非言語的コミュニケーションの欠如や技術的な障壁が新たな課題を生じさせている [4]。ハイブリッド会議は、対面とオンラインの利点を組み合わせるが、その複雑性が合意形成のプロセスに影響を及ぼす可能性

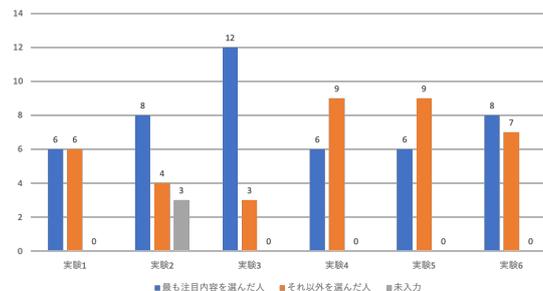


図 1.1: Enter Caption

がある。

これらの課題は、ファシリテーターの役割をより複雑にし、効果的な合意形成を困難にしている。

1.2 研究目的

本研究の目的は、オンラインおよびハイブリッド会議における合意形成のプロセスを支援するための新たな手法を開発することである。具体的には、ファシリテーターが会議中の重要な提案や意見を効率的に識別し、合意形成を促進するための支援ツールを提供することを目指す。この研究は、会議の効率化と参加者の満足度向上に寄与し、リモートワークやハイブリッドワークの普及に伴う新たな課題に対処することを意図し、オンライン及びハイブリッド会議における合意形成のプロセスを支援する新たな手法の開発に焦点を当てている。オンライン会議の普及に伴い、非言語的コミュニケーションの欠如や技術的な障壁といった新たな課題が効果的な合意形成を困難にしており、ファシリテーターの役割はより重要なものとなっている。本研究では、これらの課題を克服し、ファシリテーターの負担を軽減するために、AI エージェントを用いた新しいアプローチを探求している。具体的には、テキストベースのオンライン会議において、ファシリテーターが参加者の提案や意見を効率的に識別し、合意形成に向けた進行を支援することが挙げられる。AI エージェントが会議中のデータを分析し、重要な情報をファシリテーターに提供することにより、合意形成の過程をスムーズに進めることが期待される。また、会議の効率化と参加者の満足度向上に寄与することを目指している。リモートワークやハイブリッドワークの普及に伴う、オンライン会議の合意形成における新たな課題に対処することもこの研究の目的である。これらを踏まえて、本研究はオンライン会議における合意形成を支援するための新しいアプローチを提案し、その実現可能性と有効性を探る。

そこで、本研究ではオンライン会議における合意形成の中で、ファシリテーターが会議における提案を整理し、会議を円滑化することを支援するためのエージェントを開発する。開発されたエージェントにより、ファシリテーターの経験や公平性の違いによる合意結果への影響を軽減し、多数の提案がある会議でも合意形成に向けた円滑な進行を支援することが期待できる。

1.3 論文の構成

本論文は、オンライン会議における合意形成を支援するファシリテーター支援エージェントに関する研究を報告するものである。

1. 第1章「序論」では、オンライン会議における合意形成の重要性と、ファシリテーターの役割について概説し、本研究の動機と目的を明確にする。

2. 第2章「関連研究」では、合意形成のプロセス、ファシリテーターの影響、およびAIエージェントを利用した会議支援に関する既存の研究をレビューし、本研究の位置付けを議論する。
3. 第3章「エージェント設計」では、オンライン会議用プラットフォームの構築と、提案の優先順位付けを行う推薦エージェントの設計原理とアルゴリズムについて詳述する。図 3.2 はそれを示したのものである。

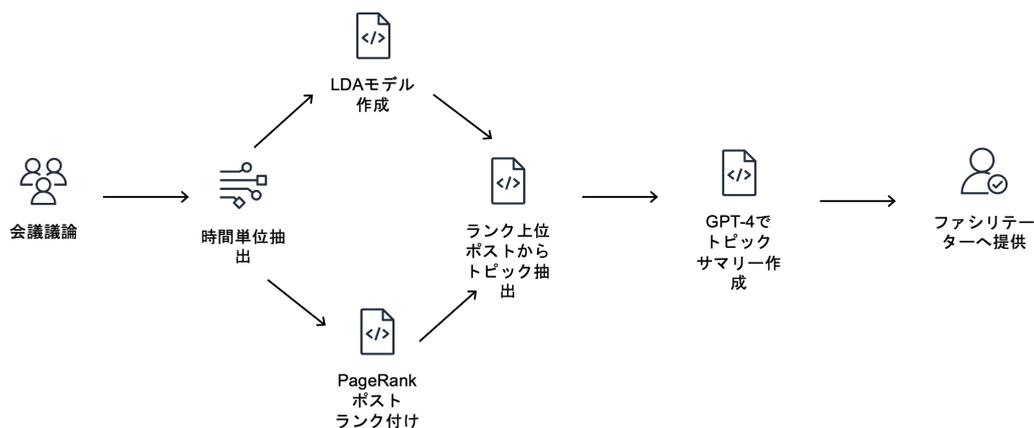


図 1.2: システムフローチャット

4. 第4章「実験」では、開発された推薦エージェントがファシリテーターをどのように支援し、合意形成にどのように寄与するかを検証するための実験設計と結果を報告する。
5. 最終章である第5章「結論とまとめ」では、研究の成果を要約し、推薦エージェントがオンライン会議の合意形成に与える影響と、今後の研究の方向性について考察する。

第2章 関連研究

第2章では、オンライン会議における合意形成とファシリテーターの役割に関連する研究を概観する。具体的には、まず、合意形成の理論的枠組みとそのプロセスについての基礎的な概念を解説する。次に、ファシリテーターが会議の進行を支援し、参加者間の意見交換を促進するために用いる手法と戦略に焦点を当てる。さらに、ファシリテーターの介入が合意形成に与える影響についての研究をレビューし、その結果がどのように会議の成果に影響を及ぼすかを考察する。

オンラインおよびハイブリッド会議の普及により、会議の進行や合意形成における新たな課題が浮き彫りになっている [5]。これには、参加者間のコミュニケーションの効率性、技術的な障壁、非言語的コミュニケーションの欠如などが含まれる。

AI技術の進歩に伴い、オンライン会議の文脈でAIエージェントが果たすことができる役割についても議論する。AIエージェントがファシリテーターの役割を補完または代替する方法、会議の効率性を高めるための推薦システムや意思決定支援システムの開発に関する最新の研究を取り上げる。このようなシステムが合意形成にどのように貢献し得るか、また、それらが持つ潜在的な限界と課題についても検討する。

2.1 会議と合意形成

会議における合意とは、議論や協議を通じて参加者が問題、提案、方針などに関して共通の理解や意見に至り、それを承認または支持することである。この共通の理解や意見に至るまでのプロセスを合意形成と呼ぶ [6]。合意形成は、参加者間のコミュニケーションと協調を必要とし、その質と効率は会議の成果に直接影響を与える。また、これは参加者間での共通理解と意見の一致を目指すプロセスであり、その達成には多くの課題が存在する。特に、オンラインの文脈では、非言語的手がかりの欠如や、技術的な障壁がこれらの課題を増大させる可能性がある。

HUらは日本語上級学習者が参加するピア・リーディング授業の一環として行われたグループ・ディスカッションにおける合意形成のプロセスを解明するために、グループ・ディスカッションの過程を詳細に分析し、発話機能の観点から学習者間の合意形成のプロセスを調査した [1]。この分析では、進行表明段階、解答提示段階、解答議論段階、解答整理段階、合意形成段階という五つの段階を特定し、意

見の分かれ方によるプロセスの変化、伝達機能の複雑化、司会進行役の影響、議論の進行ストラテジーなどを観察した。その結果、合意形成のプロセスは五つの段階を踏むこと、意見の分かれ方によってプロセスが変化すること、伝達機能が複雑化する傾向があること、司会進行役の存在が議論に影響を与えること、そして議論がまとまらない時の特定のストラテジーが有効であることが明らかになった。これらの発見は、日本語教育におけるグループ・ディスカッションの進行方法や学習者間の合意形成を促進するための指導法の開発に役立つ洞察を提供している。

2.2 オンライン会議とファシリテーター

新型コロナウイルスの影響により、リモートワーク、遠隔教育、オンライン会議などの新しいライフスタイルが加速した [7]。これに伴い、オンライン会議は地理的な制約や異なるタイムゾーンを克服し、異なる場所からの参加を可能にし、柔軟なスケジュールでの開催を実現している。コミュニケーションツールの多様性が豊かになり、テキストチャット、ビデオ会議、音声通話など、参加者が自分に適したコミュニケーション方法を選択できるようになった。これにより、意見交換がより円滑に行えるようになっている。

そのうち、テキストベースのオンライン会議は、参加者が自分の都合に合わせてメッセージを送受信できるため、異なるタイムゾーンにいる人々とも効果的にコミュニケーションを取ることができる [8]。また、図や図表と組み合わせて使用することで、複雑なアイデアやデータをより分かりやすく説明することが可能になる。さらに、参加者全員が質問や意見を送信することで、会議の進行をスムーズにし、理解度を確認できるという利点がある。

ファシリテーターは、会議やワークショップなどの集団的な討議の場において進行割を担う [9]。彼らの主な任務は、参加者の協力を引き出し、ディスカッションや活動を進行させ、目標達成に向けて支援を行うことである。ファシリテーターは、会議の内容が議題やテーマから逸脱しないように監視し、発言者や意見内容の偏りが生じないように注意を払い、もし偏りがあれば方向修正を行う。このようにして、会議の進行を円滑化し、効果的な意思決定や合意形成を促進する。

ファシリテーターが会議をコントロールする行為は、ファシリテーションと呼ばれ、参加者全員が平等に意見を述べる機会を持ち、集団の知識と経験を最大限に活用するプロセスを指す。ファシリテーターは、中立的な立場を保ちながら、会議の目的に沿った構造を提供し、参加者が自由にかつ建設的に意見を交換できる環境を整える。彼らは、議論を促進し、参加者間のコミュニケーションの障壁を取り除き、集団の思考プロセスを導くための技術や手法を駆使する。

効果的なファシリテーションにより、参加者は自身の考えを明確に表現し、他者の視点を理解し、共通の目標に向けて協力することが可能になる [10]。これは、特に意見が分かれるような複雑な問題に対して、共有の理解を築き、合意に至る

ために不可欠である。ファシリテーターの技術と能力は、会議の成果に大きく影響を及ぼし、そのためには継続的な研修と実践を通じて磨かれる。

Griffithらはグループサポートシステム（GSS）におけるファシリテーターの影響に焦点を当てた研究である [2]。研究の目的は、ファシリテーターがグループのコミュニケーションや意思決定プロセスにどのように影響を与えるか、またその影響が意図的なものと意図しないものの両方を持つことを理解することにある。研究では、ファシリテーターの役割、権力、情報処理、影響力などの概念を分析し、ファシリテーターの影響がグループにどのように及ぼされるかを検討した。特に、情報処理の側面（フレーミング、アンカリング、顕著性）を通じてどのように影響が及ぶかに焦点を当てた。研究の結果は、ファシリテーターが持つ影響力の源泉と、その影響力がグループの決定や成果にどのように影響を与える可能性があるかを明らかにした。ファシリテーターが意図せずにグループの成果に影響を与える可能性や、自身の議題を持つことがグループの成果にどのように影響を与えるかが示された。この研究は、ファシリテーターのトレーニングと基準設定が彼らの影響力を適切に管理するためにどのように役立つかについての洞察を提供したとともに、GSSの使用と組織のグループプロセスを効率化するための理解を深めるのに寄与し、ファシリテーターの役割と影響力をより適切に管理するための戦略の重要性を強調した。

この研究では、人間のファシリテーターが直面する問題には、偏見、一貫性の欠如、疲労やストレス、すべての参加者との平等な関与の困難さが挙げられている。これらの問題に対処するため、AIにより自動化されたファシリテーションエージェントは人間のファシリテーターの不足点を補い、より効率的かつ効果的なグループの支援を提供できると期待される。

Itoらは大規模なオンライン討論を効率的に進行させるための自動化されたファシリテーションエージェントを開発し、その有効性を検証している。これは、人間のファシリテーターが直面する大規模討論の難しさを解決するために行われた。それを実現するため、テキストベースの意見から討論の構造を抽出し、ファシリテーションメッセージを生成して投稿する自動化されたファシリテーションエージェントを開発した。[11] このシステムは、IBIS (Issue-Based Information System) 構造を採用し、意見、利点、欠点などを識別した。名古屋市役所との大規模な社会実験を実施し、自動化されたファシリテーションエージェントの効果を評価した。実験の結果、自動化されたファシリテーションエージェントは人間のファシリテーターと比較しても十分に機能することが示された。特に、自動化されたファシリテーションエージェントと人間のファシリテーターが協力する場合、最も高い満足度が得られた。また、自動化されたエージェントによるファシリテーションは、アイデアや利点の生成において非常に効果的であることが示された。この研究は、大規模オンライン討論のための自動化されたファシリテーションエージェントの有効性を実証し、民主的な市民プラットフォームの次世代手法としての可能性を示している。

しかし、自動化されたファシリテーションエージェントは、複雑な人間の感情や意思決定のプロセスを完全に理解および模倣することに限界がある。特に、皮肉や微妙な意見の違いを捉えるのは難しい。人間のファシリテーターに比べて、自動化されたエージェントに対する参加者の信頼や受容性が低い可能性がある。また、システムが特定のタイプの意見や議論スタイルを優先する可能性があり、多様な視点や意見が十分に反映されない恐れがある。オンライン討論のデータを収集し分析する過程で、個人のプライバシーやデータのセキュリティに関する懸念が生じる可能性がある。これらの懸念点から、現時点での自動化されたファシリテーションエージェントは人間のファシリテーターを完全に代替することはできず、人間のファシリテーターのサポートとして機能するエージェントが最適であるという結論が導かれる。

2.3 従来の会議促進エージェント

2.3.1 Kinoshita らの研究

Kinoshita らによる研究では、参加者が議論に積極的に参加し、意見を述べることを促す情報を自動的に推薦するエージェントが提案されている [12]。このエージェントは、進行中の Web ディスカッションから必要なデータを取得し、リアルタイム検索によって推薦すべき情報を決定する。検索クエリの用語は、事前に訓練されたクエリ用語生成モデルを使用して生成される。検索で得られた情報から推薦情報を選択する際には、議論のフェーズに応じて取得情報を分類するモデルが使用される。

このエージェントの介入によるディスカッション実験の結果は、エージェントの有効性を示唆するものであったが、改善点も指摘されている。特に、情報量の増加により、参加者が議論の内容を理解することは向上したが、同時にファシリテーターの負担が増加するという問題が明らかになった。ファシリテーターは、参加者からの発言を促すための情報を選択し、提供する役割を担う。しかしながら、情報の量が多くなると、どの情報が議論にとって最も関連性が高いかを判断することがより複雑になる。したがって、ファシリテーターの負担を軽減しつつ、議論の質を高めるためには、エージェントの推薦する情報の選択基準をさらに洗練させる必要がある。

2.3.2 Yoshimura らの研究

Yoshimura らは、GPT-3 を用いた話し言葉の議論の半自動要約手法を提案している [13]。研究では、フェイス・トゥ・フェイスのワークショップにおけるアイデアの発表を記録するために、音声認識技術と GPT-3 の応用を試みたものである。

研究者たちは、GPT-3を微調整し、発言の前文脈をモデル入力に加えることで、記録すべき発言の選択精度を向上させた。この手法により、F1値が0.8を超える選択精度を実現し、ROUGE-1のF1値が最大0.48に達する要約と言い換えの性能を示した。これにより、会議のオペレーターの負担を軽減することに成功している。

しかし、本研究には限界も存在する。句毎に内容を要約するアプローチでは、多数の発言者が存在する会議において、どの内容が注目すべきかの判断や、議論の全体像を把握することが困難である。会議の流れや主要テーマを正確に捉えるためには、議論の構造を理解し、発言者間の関連性を把握する高度な要約技術が求められる。したがって、今後の研究においては、会議のダイナミクスを捉えるための新たな技術開発が必要とされている。

2.3.3 Itoらの研究

Itoらは大規模オンライン議論における「フレーミング」(攻撃的なメッセージの投稿)の問題に対処し、より効果的な議論を促進するためのシステムの開発を目指し、オンライン議論を適切に導くためのファシリテーターの役割の強化とその影響について調査している [14]。それを実現するため、「COLLAGREE」というファシリテーター支援機能を持つインテリジェントな群衆意思決定支援システムを開発し、複数の実世界オンライン討論サポートのケーススタディで展開した。討論の質を高めるための様々な機能(議論ポイントシステム、議論ツリー、品質評価メカニズム)を導入した。その後、名古屋市、愛知県、国際会議などで実際のフィールド社会実験を実施し、システムの有効性を評価した。

その結果、討論の質向上とフレーミングの減少が観察された。特に、ファシリテーターの社会的存在感が参加者の行動に大きな影響を与えることが示唆された。複数のケーススタディで、システムを使用した討論においてフレーミング現象が一切観察されなかった。今後の研究の方向として、ファシリテーターの社会的存在感の効果をより制御された実験で調査すること、および自動化されたファシリテーター(人工知能エージェント)の開発が挙げられている。この論文は、大規模オンライン討論におけるファシリテーターの役割と影響に焦点を当て、実際の社会実験を通じてその効果を検証する。

2.3.4 Okuharaらの研究

Okuharaらはビデオ会議におけるエージェントの介入が人々の意見変更にも与える影響を分析することを目的としている [15]。具体的には、エージェントが参加者と反対の意見を表明し、この介入がディスカッションの流れや参加者の意見にどのような変化をもたらすかを検証している。研究では、エージェントの介入があるグループとないグループに分けられた14人の被験者に対して実験が行われた。

エージェントは参加者と異なる意見を表明する設定であり、その介入が議論にどのような影響を与えるかが観察された。

実験の結果、エージェントの介入があるグループでは意見変更が見られたが、統計的に有意な効果を明確に示すには至らなかった。これは、参加者数が少ないため、統計的な検証に適していなかったと考えられる。そのため、将来的にはより多くの参加者を対象にした実験が求められている。この研究は、ビデオ会議やオンラインディスカッションにおけるエージェントの効果的な介入方法に関する理解を深めることを目指しているが、追加実験が必要であることを示唆している。

2.3.5 Yangらの研究

Yangらは、インターネットの発展に伴い、オンライン議論システムや社会民主主義システムがグループ意思決定支援のための重要かつ効果的な手段となっていることを背景に、大規模なオンライン議論においてコンセンサスに到達するための機械学習をベースとしたコンセンサス決定支援システム (CDMS) を提案した [16]。既存のオンライン議論システムでは、経験豊富な人間のファシリテーターが議論をナビゲートし、ガイドする必要があるが、これには多くの課題がある。そこで、インテリジェントなファシリテーションの探求が必要不可欠であり、本研究はそのための方法を提案した。研究では、他者と異なる意見を表現するエージェントを実装し、その効果を検証することで、オンライン議論におけるエージェントの介入効果を調査した。COLLAGREE というオンライン議論プラットフォームを用いて、開発した機械学習ベースの方法をオフラインで評価し、初期の実験結果は、開発された機械学習ベースの方法が CDMS に有用であることを示し、その実現可能性を示した。

この研究から、オンライン議論や集団意思決定支援のためのインテリジェントなファシリテーション手法として、機械学習ベースのアプローチが有効であることが明らかになった。しかし、より高度な CDMS 機能の開発や異なるドメインへの適用など、さらなる研究が必要であることも示された。

2.3.6 Fujitaらの研究

Fujitaらは、日本とアメリカの共同プロジェクトにおいて、オンラインでの大規模議論を支援する先進的なコンピューターシステムを開発していることを説明している [17]。これは、より賢く、よりつながったコミュニティを実現するために必要とされる。共同作業では、大規模なオンライン議論でしばしば発生する2つの問題、つまり、(1) アイディエーション：群衆がより効果的に勝ち勝ちの解決策を開発するのを助け、(2) 意思決定：群衆が選択した解決策においてパレート最適性に到達するのを助ける、に取り組むことに焦点を当てている。日本では、

自由なテキスト議論を促進してコンセンサスを達成するための議論支援システム「COLLAGREE」が開発されている。一方、アメリカでは、議論理論とソーシャルコンピューティング技術を統合して、より効果的な大規模議論を可能にするオンラインツール「Deliberatorium」が開発されている。直近の共同作業の1つとして、COLLAGREEの自由テキスト議論とDeliberatoriumによって提供される構造化された議論を統合することが挙げられる。また、より良いアイディエーションと意思決定を可能にする自動化エージェントの開発も行われる。

この研究から、COLLAGREEとDeliberatoriumのようなシステムを統合することで、参加者がより効果的にアイデアを生成し、選択した解決策においてパレート最適性に近づくのを助けることができることが示された。これは、大規模なオンライン議論において賢く、つながったコミュニティを実現するために必要な支援を提供することを目指している。

2.4 本研究の位置付け

本研究は、ファシリテーターが議論の進行を理解するための情報量を最適化することに重点を置いている点にある。従来の研究では、オンライン会議やディスカッションにおいて、ファシリテーターが参加者からの発言を促すための情報を選択し、提供する役割を担っていた。しかし、情報の量が多くなると、どの情報が議論にとって最も関連性が高いかを判断することが複雑になり、ファシリテーターの負担が増加するという問題が明らかである。本研究では、この問題に対処するために、情報量を効果的に管理し、ファシリテーターの負担を軽減することを目指している。

期待される効果としては、合意形成の過程をスムーズにすることが挙げられる。提案された手法により、ファシリテーターは会議の流れをより効率的に管理し、重要なトピックや議論ポイントに焦点を当てることができる。これにより、参加者は自身の意見をより明確に表現し、他者の視点を理解しやすくなり、共通の目標に向けて協力しやすくなる。また、ファシリテーターの負担が軽減されることで、会議の進行がスムーズになり、参加者の満足度が向上することが期待される。

第3章 ファシリテーター支援エージェント

本章では、オンライン会議における、ファシリテーターが合意形成を支援するエージェントを提案する。

3.1 オンライン会議の構造

オンライン会議の構造は一般に図 3.1 に示すように、「ポスト」、「トピック」、「返信」、「いいね」要素を含む。ポストは参加者が投稿する個々のメッセージやコメントであり、議論の内容を形成する。トピックは議論の主題やテーマを指し、会議の焦点となる中心的な概念である。返信は他の参加者のポストに対する直接的な反応やコメントであり、議論を深めるために用いられる。いいねは、特定のポストや返信に対する肯定的な反応を示す機能であり、参加者の支持を表現する。これらの要素は、オンライン会議の流れや参加者間の相互作用を形成し、議論の進行に重要な役割を果たす。

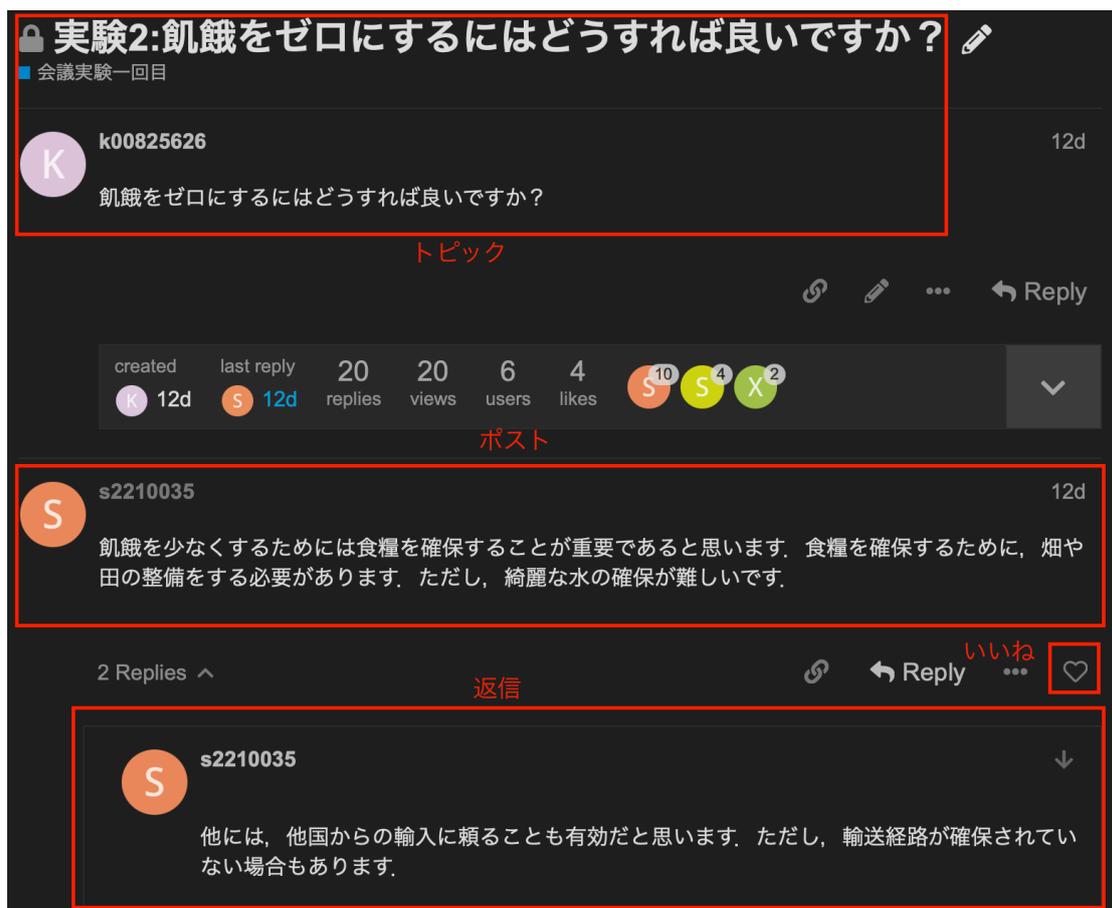


図 3.1: オンライン会議の構造

3.2 エージェントの役割

エージェントの役割として、テキストベースのオンライン会議において、参加者たちの議論中から、より多くの人々が注目した内容をファシリテーターへ提供することで、ファシリテーターの情報処理の負担を軽減し、参加者間の議論の促進や意見交換の円滑化に集中できるようにすることが挙げられる。これを実現するためには、トピックの抽出と注目度の判別を実現することが不可欠である。

3.3 エージェント処理の流れ

1. 会議議論の収集: オンライン会議プラットフォームからテキストベースの議論データを抽出する。
2. LDA モデル [18][19] による注目内容の抽出: 抽出されたテキストデータに LDA モデルを適用し、会議の中で議論された主要なトピックや話題を抽出

する。このプロセスにより、議論の中心となる内容やトピックが明らかになり、それぞれのポストがどのトピックに当てはまるかが分かる。

3. PageRank[20] による注目順位付け: 会議議論のポストの返信関係を使い、PageRank アルゴリズムを適用し、それぞれのポストの注目度を評価する。このステップでは、ポスト間の関連性や参照の頻度に基づいて、ポストの注目度を PageRank 値で表示する。
4. 注目トピックの決定:LDA モデルは会議全体の発言により生成される。そして、注目ポストの判定は、指定時間内のポストの PageRank による評価結果を基に、この時間帯で最も注目されたポストを特定する。その後、LDA モデルを使って、注目ポストがどのトピックに当てはまるかを判別し、トピックの確率分布が最も高いものを注目トピックに決定する。これにより、ファシリテーターは会議の焦点を絞り、効率的な議論進行と合意形成を支援できるようにことが期待される。

図 3.2 はこの流れを示したものである。時間単位で抽出された議論内容に対して、トピック抽出処理とポスト注目度ランク付けに分かれて処理が行われ、それらを合流させて注目トピックの判別を行う。次節ではそれぞれの役割について説明する。

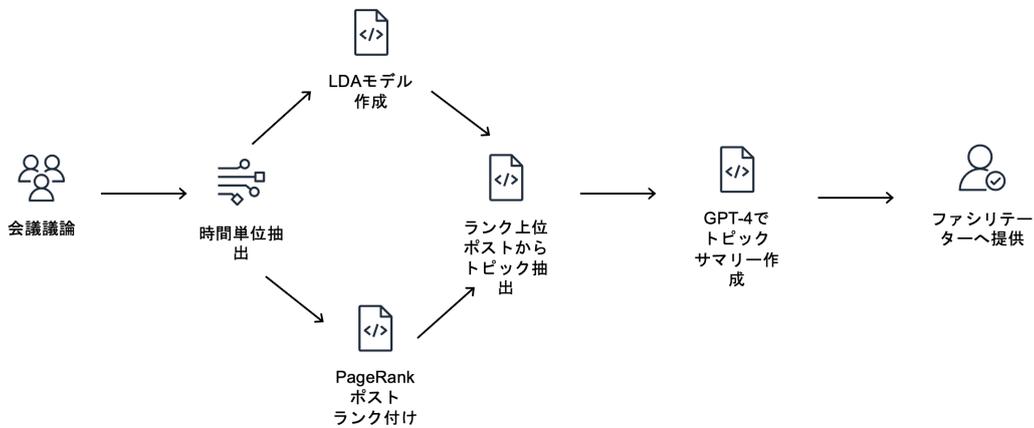


図 3.2: システムフローチャット

3.4 トピック抽出

トピック抽出の仕組みは、テキストデータから特定のトピックやテーマを識別し、分類するプロセスである。このプロセスは、大量の文書やテキストデータに含まれるパターンや傾向を解析し、それらを意味のあるカテゴリーに分けることを目的としている。通常、トピック抽出は自然言語処理（NLP）技術を用いて行われ、特に機械学習のアルゴリズムが広く利用される。本研究でトピック抽出する

手法としては、Latent Dirichlet Allocation (LDA) [18][19] を使う。この手法は、文書集合を入力として受け取り、各文書がどのトピックに関連しているかを確率的に推定する。例えば、LDA は文書が複数のトピックの混合で構成されていると仮定し、各トピックが特定の単語の集合によって表現されると考える。このモデルは、文書内の各単語がどのトピックに属するかを確率的に割り当て、その結果を基に文書全体のトピック分布を推定する。

トピック抽出のプロセスは、まず文書集合を前処理することから始まる。この段階では、テキストの正規化、単語のトークン化、ストップワードの除去、語幹の抽出などが行われる。次に、選択したトピックモデルを用いて文書集合を分析し、各文書に含まれるトピックを識別する。最終的に、モデルは各トピックに関連する単語のリストと、各文書がどのトピックにどの程度関連しているかを示すトピック分布を出力する。この出力に基づいて、研究者や分析者は文書集合内の主要なテーマやトピックを理解し、さらなる分析や意思決定のための洞察を得ることができる。

3.4.1 LDA モデルの概要

LDA モデルは、文書が複数のトピックから成り立っていると仮定する。ここでの「トピック」とは、特定の単語群の集合を指し、それぞれの文書はこれらのトピックの確率的な混合によって生成されると考え、各トピックの割合をトピックの確率分布とする。LDA の目的は、与えられた文書集合からこれらのトピックを推定することである。

図 3.3[21] は、LDA モデルを示したものである。左側に示された「Topics」には、トピックごとに関連する単語とその重みが表示されている。中央の大きな図は、ある文書がどのように異なるトピックから構成されているかを示している。文書内の各単語は、色分けされたノードを通じて特定のトピックに割り当てられる。右側の「Topic proportions and assignments」は、文書におけるトピックの割合を棒グラフで視覚的に表示している。

このモデルでは、トピックの割り当てとその比率を通じて、文書がトピックの混合からどのように生成されるかを学習する。LDA は教師なし学習アルゴリズムの一種であり、特定のトピック数を事前に設定する必要がある。学習プロセスが完了すると、新たな文書が与えられた際に、その文書が含むトピックの割合を推定することができる。

LDA の応用例としては、文書分類、コンテンツ推薦、検索システムなどがある。トピックモデルリングによって、大量のテキストデータから有用な情報を抽出し、データの構造を理解するのに役立つ。

例に、5つの単語で構成される2つのトピックを抽出した結果を図 3.5 に示す。それぞれのトピックは単語によると、政治的議論 (Political debates) とサッカー (Football) のトピックである。トピックについての解釈は、人それぞれの意見によって違うが、トピックが存在していることがわかる。

```
documents = [  
  "I enjoy watching football matches",  
  "Political debates are usually intense",  
  "Football is a great sport"  
]
```

図 3.4: LDA 処理文の例

```
(0, '0.124*"usually" + 0.124*"intense" + 0.124*"debates" + 0.124*"political"')  
(1, '0.147*"football" + 0.088*"great" + 0.088*"a" + 0.088*"sport"')
```

図 3.5: 例文の処理結果

3.5 ポスト注目度算出

ポスト注目度の算出において、PageRank アルゴリズムの使用は、オンライン会議やディスカッションフォーラムにおける議論の中から特に注目されるべきポストを特定する効果的な手段である。

3.5.1 PageRank とは

PageRank とはもともとウェブページの重要性を評価するために開発されたアルゴリズムであり、ウェブページ間のリンク構造を分析して各ページの重要度を数値化する [20]。このアルゴリズムを議論の文脈に応用することで、各ポストの重要度を評価し、議論全体の中でのその立ち位置を明確にすることができる。

具体的には、会議やディスカッション中の各ポストをノードと見なし、ポスト間の返信関係をエッジとしてネットワークを構築する。PageRank アルゴリズムをこのネットワークに適用することで、各ポストの重要度を数値化し、PageRank 値が高いポストを注目ポストとして特定することができる。このプロセスでは、ポスト間の相互関係が重要度の評価に直接影響を与えると仮定し、PageRank アルゴリズムを使用することで、どの発言が会議全体において最も影響力があるかを定量的に評価する。

この方法により、ファシリテーターは会議やディスカッションの中で特に重要なポイントや話題を効率的に特定できる。また、PageRank 値に基づいて注目ポストを抽出することで、会議の合意形成において重要な役割を果たすポストを明確

にし、議論の進行を効果的にサポートすることが可能となる。このアプローチは、特に情報量が多い会議や複数の議題が交錯する複雑な議論において、その価値を発揮する。

3.5.2 PageRank アルゴリズムの説明

PageRank アルゴリズムの式は以下の通りである。

$$PR(A) = (1 - d) + d \sum_{i=1}^n \frac{PR(T_i)}{C(T_i)}$$

- $PR(A)$ はノード A の PageRank 値
- d は減衰係数であり、通常 0.85 に設定される。これは、ランダムなユーザーが任意のページからリンクをたどって別のページに移動する確率を表す。
- $1 - d$ は閲覧者がリンクをたどらずにネットワーク内の任意のノードにランダムにジャンプする確率を表している
- T_i はノード A に向けられた全ノードの集合の中の一つのノード
- $PR(T_i)$ はノード T_i の PageRank 値
- $C(T_i)$ はノード T_i のアウトリンク数、つまり T_i から他のノードへ向けられたリンクの数である
- $\sum_{i=1}^n \frac{PR(T_i)}{C(T_i)}$ は A に向けられた全ノードの PageRank 値をそれぞれのアウトリンク数で割った合計

オンライン会議の PageRank 値を計算する場合は、会議の記録からネットワークを構築し、各発言はノードとして表現され、発言間の相互作用（例えば、一つの発言が別の発言に返信している場合）はエッジとして表現される。そして重要性を計算する場合は、PageRank アルゴリズムは、ネットワーク内の各ノード（発言）に対して重要性のスコアを割り当てる。このスコアは、他のノード（発言）からの「支持」の数と質に基づいており、ノードへのリンク（返信や言及）は他のノードからの支持とみなされる。そして減衰係数 d （通常は 0.85 に設定）は、参加者が現在の話題や議論にとどまり続ける（つまり、関連する返信や発信を続ける）確率を表す。ランダムジャンプ $1 - d$ の部分は、参加者が話題から離れ、新しい話題に移行するか、会議から離脱する確率を示す。

実際の応用では、PageRank アルゴリズムはネットワーク内のすべてのノード上で繰り返し計算され、各ノードの PageRank 値が安定するまでこのプロセスが続く。これにより、最も重要あるいは影響力のあるノードを特定できる。

図 3.6 はポストの構造を示した一例である。ポスト 1 が最初の投稿で、ポスト 2 とポスト 3 がポスト 1 に対する返信、ポスト 4 がポスト 2 に対する返信であり、ポスト 5 はポスト 4 に対数 r 返信である。

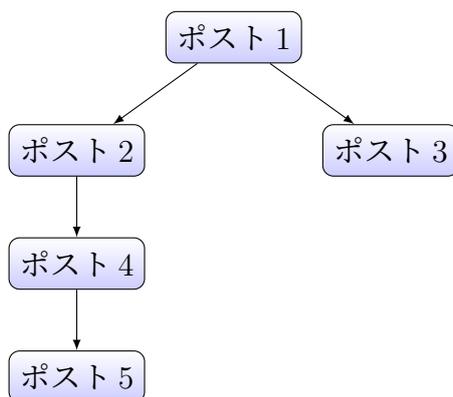


図 3.6: オンライン会議におけるポストの構造例

減衰係数 0.85 を設定し、PageRank を計算を行ったところ、それぞれのポストのランキングは以下の通りとなった。ポスト 2 とポスト 3 はいずれもポスト 1 に対する返信であるが、さらに返信が続くポスト 2 の方が、ポスト 3 よりも PageRank は大きな値となる。この場合は、ポスト 2 がポスト 3 より注目されていることを意味する。

1. ポスト 1 の PageRank 値: 0.3859
2. ポスト 2 の PageRank 値: 0.2460
3. ポスト 4 の PageRank 値: 0.1769
4. ポスト 3 の PageRank 値: 0.0956, ポスト 5 の PageRank 値: 0.0956

3.6 注目内容のトピック判別

LDA モデルによるトピック判別と PageRank による注目ポストの判断を組み合わせることで、会議における注目トピックをより総合的に特定することが可能となる。LDA モデルが明らかにしたトピックの分布を使って、PageRank アルゴリズムによって特定された注目ポストに含まれるトピック確率分布が最も大きいトピックを注目トピックとして、GPT-4 でトピックのサマリーを作成する。これにより、会議の中で特に重要なテーマや議論のポイントを包括的に把握することができる。つまり、会議の中でどのトピックが最も関心を集めているか、またそのトピックに関連する重要な発言は何かを明確にすることが可能となり、ファシリテーターはより効果的に会議を進行し、合意形成を促進するための具体的なアクションを取ることができる。

GPT (Generative Pre-trained Transformer) は、自然言語処理 (NLP) における深層学習モデルの一種で、OpenAI によって開発された。このモデルは、Transformer アーキテクチャ [22][23] に基づいており、大規模なデータセットを用いた事前訓練によって、広範な言語理解と生成能力を獲得する。Transformer は、注意機構

(Attention Mechanism) を中心に構築されている。この機構により、モデルは入力されたテキストの各部分に対して、どの程度の「注意」を払うべきかを学習する。これにより、文脈の理解が深まり、より関連性の高いテキスト生成が可能になる。

GPT モデルは、大量のテキストデータを用いて事前に訓練される。この事前訓練プロセスでは、モデルは言語の基本的なパターンや構造を学習する。その後、特定のタスクに合わせて微調整 (Fine-tuning) が行われる。これにより、モデルは特定の応用分野においても高い性能を発揮する。バージョンとして、GPT-1 から始まり、GPT-2、GPT-3 と進化を遂げてきた。各バージョンは、モデルのサイズ (パラメータの数)、文脈処理能力、多言語対応能力などの点で改善が図られている。そのうち、GPT-4 は、現時点 OpenAI によって開発された自然言語処理モデルの最新版である [24]。このモデルは、前身である GPT-3 の基本構造を踏襲しつつ、より大きなモデルサイズとパラメータ数を持ち、これによりより複雑な文脈を理解し、精度の高いテキスト生成が可能である。GPT-4 は特に長い文脈を処理する能力が向上しており、長い会話や文書においても一貫性のある応答を提供する。

PageRank で評価された注目度の高いポストのトピックから抽出された単語を利用して、GPT モデルを用いてトピックのサマリを作成する。サマリは、会議の主要な議論ポイントを簡潔に伝え、参加者が議論の全体像を把握するのに役立つ。また、会議の後のレビューやフォローアップの際にも、重要なトピックに素早くアクセスし、深く理解するための基盤となる。

このように、PageRank で評価された注目ポストのトピックから単語を抽出し、それを基に GPT モデルを用いてトピックのサマリを作成する方法は、オンライン会議の効率化と理解の深化を促進する強力なツールとなる。また、これらのプロセスにより、会議の重要な内容を迅速かつ正確に要約し、参加者が議論の本質を捉えることを支援することが可能である。つまり、オンライン会議のデータ分析の精度と有用性を高めることができ、会議の効率化と参加者の満足度向上に寄与することが期待できる。

第4章 実験と考察

本章では、提案された手法の有効性を検証するための実験と、その結果に基づく考察を行う。

4.1 実験概要

4.1.1 実験目的

実験の目的は、オンライン会議における合意形成を支援するために開発されたエージェントの効果を評価することである。具体的には実験の参加者を集め、オンラインプラットフォームで会議実験を行う。その後、実験のデータを使って、LDAモデルを作成し、PageRankで注目ポストを特定する。最後に、LDAモデルで、注目ポストにおいてトピックの確率分布が最も高いものを注目トピックとする。さらに、GPT-4でサマリー生成後、参加者のアンケートと比較して提案手法の妥当性を評価する。また、実際の会議データに基づく分析により、提案手法の実用性と限界を明らかにし、今後の改善点を特定する。

なお、本実験は北陸先端科学技術大学院大学ライフサイエンス委員会(人05-061)の承認を受けたものである。

4.1.2 実験対象

オンライン議論プラットフォームにおいてテキストベースの議論を行うため、日本語会話能力を持ち、一定のキーボード入力経験を持つと考えられる本学の学生を対象者とした。参加者はそれぞれ、異なる背景を持つ個人であり、多様な視点から議論に貢献することが期待される。

4.2 実験方法

4.2.1 実験環境

実験の条件をなるべく同一にするため、参加者は会議室に集合し、各自で持ち込んだノートパソコンで会議実験を行った。実験は同じ日の午前と午後に計6回

実施した。1回の実験は5名で構成されたが、一部の実験は参加者が欠席したため4名で議論した。実験の参加者は合計8名であり、2名は午前と午後の実験両方参加した。

4.2.2 実験の流れ

1. 会議の進行: 各グループは、オンラインプラットフォームを使用して30分間の会議を行う。会議のテーマは事前に決定され、参加者はこのテーマに沿って自由に議論を進める。
2. 注目内容の記録: 会議中、10分ごとに参加者に対して、その時点での会議の注目内容を記録してもらう。これにより、会議の進行に伴う注目点の変化を捉えることができる。
3. データ収集: 会議の全テキストデータを収集し、提案手法を用いて分析を行う。分析には、LDAモデルによる注目内容の抽出、PageRankによる優先順位付けが含まれる。
4. 評価方法: 実験終了後、参加者による注目内容の記録と提案手法による分析結果を比較評価する。この比較により、提案手法の精度と実際の参加者の認識との一致度を評価する。

4.2.3 実験アンケート

実験アンケートは、参加者からの注目内容を選択するために用いられる。このアンケートは、オンライン議論の各段階で参加者がどのトピックや議論ポイントに最も関心を持ったかを特定することを目的としている。アンケートには、議論中に取り上げられた主要なトピックや議論ポイントがリストアップされ、参加者はそれらの中から最も注目した内容を選択する。このプロセスにより、参加者の視点から見た議論の重要な側面が明らかになり、提案された手法による分析結果との比較が可能となる。

4.3 実験システム

4.3.1 オンライン会議実験用プラットフォーム

オンライン会議の実験用プラットフォームは、オープンソースの「Discourse Meta」を使用して構築した。このプラットフォームは、ConoHaで提供されるサーバー上に設置され、no-ipを通じて無料のホスト名が利用されている。Discourse Metaは、柔軟なカスタマイズが可能なディスカッションフォーラムソフトウェアであり、オンラインでの議論やコミュニケーションに適している。ConoHaのサーバー

表 4.1: 実験用プラットフォーム

OS	Ubuntu18.0
Open source platform	Discourse meta
SSD	100GB
CPU	2 Core
メモリ	1GB
ネットサイト	https://disme.hopto.org

表 4.2: 実行必要なライブラリのバージョン

Python	3.11.6
mecab-python3	1.08
mecab-ipadic-neologd	2020-08-20
gensim	4.3.2

は安定した運用と必要なリソースを提供し、no-ip によるホスト名はアクセスの容易さを保証する。この組み合わせにより、実験参加者は安定した環境でオンライン議論を行うことができ、実験のデータ収集と分析が効率的に行われる。表 4.1 は実験用プラットフォームの情報を示したものである。表 4.2 は実行で使われたライブラリのバージョンを示したものである。

4.3.2 トピック抽出機能

トピック抽出するには、Python 言語を使用して gensim ライブラリーおよび MeCab ライブラリーを活用し、オンラインフォーラムのデータからを使って LDA モデルを作成してから、トピックを抽出する。具体的な役割を以下に示す。

- データ取得: Discourse API を利用して特定のトピックからポストを取得する。また、利用制限を回避するため、API キーとユーザー名を用いて認証し、特定のトピック ID に対するポストを JSON 形式で取得する。
- データクリーニング: 取得したポストの HTML をクリーニングするために `clean_html` 関数を使用する。正規表現と HTML アンエスケープ処理を通じて、テキストから HTML タグを除去する。
- 形態素解析: MeCab を使用して、クリーニングされたテキストから有意な単語（名詞、動詞、形容詞）を抽出する。ストップワードリストを利用して、必要に応じて単語の添付や削除を行い、一般的で無意味な単語を除外する。ここで、よりいい形態素解析結果を求めるため、MeCab 用の日本語形態素解析器の辞書 "mecab-ipadic-NEologd" を使った。これは、標準の MeCab 辞書 (IPAdic) に基づいているが、新語、俗語、インターネットスラング、固有名詞などのエントリーを大幅に拡張している。

- トピックモデリング:gensim ライブラリーの LDA (Latent Dirichlet Allocation) モデルを使用して、文書からトピックを抽出する。トピック数、更新回数、パス数などのパラメータを設定し、モデルをトレーニングする。
- モデル評価:LDA モデル一貫性スコア (coherence) を計算し、モデルの品質を評価する。一貫性はモデルが抽出したトピックがどれだけ意味的に一貫しているかを測る指標で、スコアは高いほど良いとされる。

アルゴリズム 4.3.1 は計算のアルゴリズムを示したものである。

アルゴリズム 4.3.1: Discourse フォーラムトピック分析

```

1: procedure ANALYZETOPIC(forum_url, topic_id)
2:   data ← FETCHTOPICDATA(forum_url, topic_id)
3:   keywords ← EXTRACTKEYWORDSFROMDATA(data)
4:   lda_model ← PERFORMTOPICMODELING(keywords)
5:   DISPLAYRESULTS(lda_model)
6: end procedure
7: function FETCHTOPICDATA(url, id)
8:   return Discourse API からトピックデータを取得
9: end function
10: function EXTRACTKEYWORDSFROMDATA(data)
11:   return データからキーワードを抽出
12: end function
13: function PERFORMTOPICMODELING(keywords)
14:   return キーワードに基づく LDA トピックモデリング
15: end function
16: function DISPLAYRESULTS(model)
17:   return モデル結果の表示
18: end function

```

4.3.3 ポスト注目度算出機能

Python 言語を使用して、Discourse フォーラムの範囲内のポスト注目度を算出するために、ポスト間の返信関係を使い、有向グラフを作成した。その後 NetworkX ライブラリーを用いて、作成したグラフに基づき各ポストの PageRank 値を計算する。PageRank 値が高い三つのポストを注目ポストとして利用する。アルゴリズム 4.3.2 は計算アルゴリズムを示したものである。

アルゴリズム 4.3.2: PageRank 計算および上位 3 位の出力

```
1: procedure CALCULATEPAGERANK(topic_id, api_key, api_username)
2:   posts ← GET_POSTS_FROM_DISCOURSE(topic_id, api_key, api_username)
3:   G ← CREATE_GRAPH_FROM_POSTS(posts)
4:   pagerank ← CALCULATE_PAGERANK(G)
5:   return pagerank
6: end procedure
7: procedure PRINTTOPTHREEPOSTS(pagerank)
8:   sorted_posts ← SORT_POSTS_BY_PAGERANK(pagerank)
9:   for i ← 1 to 3 do
10:    PRINT sorted_posts[i]
11:  end for
12: end procedure
13: result ← CALCULATEPAGERANK(post_id, 'your_api_key', 'your_username')
14: PRINTTOPTHREEPOSTS(result)
```

4.3.4 注目ポストのトピック判別と LDA モデルに基づくサマリー生成

Python 言語を使用して、4.3.2 節と 4.3.3 節で算出した LDA モデルと PageRank 上位 3 位のポストの中からトピックの確率分布が最も高いトピックを注目トピックとする。そして、openai の GPT-4 の API を利用して入力し、設計したプロンプトに基づいてサマリーを出力し、プロンプトは以下で示している。

- 以前の内容を見捨てる
- あなたは会議のアシスタント
- 会議内容を使って、LDA モデルを作った。
- LDA モデルはトピックを出力し、トピックは単語の塊
- 上下文関係とトピックの単語に従って、できるだけトピック単語を利用し、トピックのサマリーを簡潔な一言で出力
- これがいい例"教育資源の不足という問題についてデジタル技術の応用 (AI の整備)"
- 説明は必要ない、それぞれトピックのサマリーだけ出力

それぞれの役割として、以前の内容に影響されないように、役割を決める、背景説明、出力内容指定、サンプルによる出力精度向上、注意事項が含まれている。アルゴリズム 4.3.3 は計算アルゴリズムを示したものである。

アルゴリズム 4.3.3: LDA モデルと PageRank に基づく会議内容のサマリー生成

```
1: procedure GENERATESUMMARY(forum_data)
2:   lda_model ← CREATELDAMODEL(forum_data)
3:   top_posts ← GETTOPPAGERANKPOSTS(forum_data)
4:   highlighted_topic ← IDENTIFYMAINTOPIC(lda_model, top_posts)
5:   summary ← CREATESUMMARYWITHGPT(highlighted_topic)
6:   PRINT(summary)
7: end procedure
8: function CREATELDAMODEL(data)
9:   return LDA モデルによるトピック分析
10: end function
11: function GETTOPPAGERANKPOSTS(data)
12:   return PageRank の上位 3 位のポスト
13: end function
14: function IDENTIFYMAINTOPIC(lda_model, posts)
15:   return トピック確率分布が最も高いトピック
16: end function
17: function CREATESUMMARYWITHGPT(topic)
18:   response ← OpenAI API を使ってサマリー生成
19:   return response
20: end function
```

4.4 実験結果

実験では持続可能な開発目標 (SDGs) 内の問題を選択し、それぞれ議論をさせた。午前の議論テーマは時間順番で

- 実験 1: "貧困をなくすにはどうすればいいですか?"
- 実験 2: "飢餓をゼロにするにはどうすれば良いですか?"
- 実験 3: "ジェンダー平等を実現するにはどうすれば良いですか?"

そして午後の議論テーマは時間順番で

- 実験 4: "質の高い教育をみんなに届くにはどうすればいいですか?"
- 実験 5: "人や国の不平等を無くすにはどうすればいいですか?"
- 実験 6: "気候変動に具体的な対策はなにかありますか?"

となる。

表 4.3: 実験データ

	実験 1	実験 2	実験 3	実験 4	実験 5	実験 6
参加者数	4	5	5	5	5	5
ポスト総数	27	49	58	28	29	33
新規ポスト	9	15	21	12	10	11
返信数	18	34	37	16	19	22

4.4.1 基礎データ

それぞれの実験のポスト数や返信数を表 4.3 に示す。

表 4.3 より、午前と午後のポスト総数は実験の進行と共に発言数は増える傾向がある。参加者が実験に慣れているかと考えている。

4.4.2 トピック抽出機能評価

LDA モデルの評価方法として、Coherence (一貫性) の値は以下のパラメータに影響される。本研究では、それらのパラメータに対し、いくつかの値を設定し、繰り返し計算を行い、最適な一貫性の値を計算する。

- *topic_nums*: トピックの数:LDA モデルによって特定されるトピックの数を指定する。この数値はデータセットや分析の目的によって異なる。実験では [3, 5, 7] を指定し最適な一貫性のスコアを計算する。
- *passes_list*: パスの数: コーパス全体に対して LDA アルゴリズムが実行する総反復回数を指定する。より多くのパスを使用すると、モデルがデータに対してより良く適合する可能性があるが、計算時間が長くなる。実験では [5, 10, 15, 20] を指定し最適な一貫性のスコアを計算する。
- *chunksize_list*: チャンクサイズ: 一度にトレーニングされるドキュメントの数を指定する。この値が大きいほど、一度の反復でより多くのドキュメントが処理され、トレーニングの速度が速くなる。実験では [5, 10, 15, 20] を指定し最適な一貫性のスコアを計算する。
- *alpha_list*: アルファトピック: 'auto' は自動的に最適な値を決定し、'symmetric' はすべてのトピックに均等な重みを与えることを意味する。実験では両者を使って、最適な一貫性のスコアを計算する。
- *update_every_list*: 更新頻度: モデルのパラメータが更新される頻度を指定する。値が大きいほど頻繁に更新され、より迅速に収束する可能性がある。ここでは [2, 1, 0] を指定し最適な一貫性のスコアを計算する。
- *random_state*: 乱数シード乱数生成器のシードを指定する。この値を固定することで、モデルのトレーニング結果の再現性を保証できる。実験を再現するため、ここで 100 と固定する。

30分のそれぞれの実験に対して、10分単位でトピックを計算した結果は以下の通りである。

表 4.4: 実験 1 の最適な LDA モデルパラメータ

	実験 1-1	実験 1-2	実験 1-3
$topic_{nums}$	3	3	5
$passes_{ist}$	15	10	20
$chunksize_{ist}$	5	10	10
$alpha_{ist}$	auto	auto	sym
$update_{everyist}$	2	1	2
$random_{state}$	100	100	100
Best Coherence (小数点以下三位)	0.658	0.666	0.552

表 4.5: 実験 2 の最適な LDA モデルパラメータ

	実験 2-1	実験 2-2	実験 2-3
$topic_{nums}$	5	3	3
$passes_{ist}$	5	5	5
$chunksize_{ist}$	5	10	20
$alpha_{ist}$	sym	sym	auto
$update_{everyist}$	0	0	2
$random_{state}$	100	100	100
Best Coherence (小数点以下三位)	0.627	0.460	0.513

表 4.6: 実験 3 の最適な LDA モデルパラメータ

	実験 3-1	実験 3-2	実験 3-3
$topic_{nums}$	7	3	3
$passes_{ist}$	15	15	5
$chunksize_{ist}$	5	20	15
$alpha_{ist}$	auto	sym	sym
$update_{everyist}$	2	2	0
$random_{state}$	100	100	100
Best Coherence (小数点以下三位)	0.576	0.483	0.490

表 4.7: 実験 4 の最適な LDA モデルパラメータ

	実験 4-1	実験 4-2	実験 4-3
$topic_{num}s$	5	7	7
$passes_{ist}$	15	20	5
$chunksize_{ist}$	5	5	5
$alpha_{ist}$	auto	auto	sym
$update_{everyist}$	2	2	0
$random_{state}$	100	100	100
Best Coherence (小数点以下三位)	0.477	0.547	0.489

表 4.8: 実験 5 の最適な LDA モデルパラメータ

	実験 5-1	実験 5-2	実験 5-3
$topic_{num}s$	3	7	7
$passes_{ist}$	5	5	15
$chunksize_{ist}$	5	20	15
$alpha_{ist}$	auto	auto	sym
$update_{everyist}$	1	2	1
$random_{state}$	100	100	100
Best Coherence (小数点以下三位)	0.662	0.620	0.603

表 4.9: 実験 6 の最適な LDA モデルパラメータ

	実験 6-1	実験 6-2	実験 6-3
$topic_{num}s$	7	3	5
$passes_{ist}$	5	5	5
$chunksize_{ist}$	5	10	5
$alpha_{ist}$	sym	auto	sym
$update_{everyist}$	0	0	2
$random_{state}$	100	100	100
Best Coherence (小数点以下三位)	0.458	0.501	0.548

このように、それぞれの実験の最適な LDA モデルのパラメータは一貫性のスコアの高さにより、選定された。これら実験結果からみると、一貫性のスコアは 0.45 から 0.67 である。

しかし、一貫性のスコアが高ければ必ず良いトピックであることを意味するわけではない [25]。人それぞれ、トピックに対する考えも異なるため、具体的に効果があるかどうかは、実験で評価しなければならない。

4.4.3 ポスト注目度算出機能評価

指定範囲内のポストの PageRank で計算したスコアが Top3 位のポストを注目ポストとする。そして、4.4.2 節で計算した、それぞれの最適な LDA モデルで Top3 位のポストに対し、その中でのトピック確率分布が最も高いトピックを注目トピックとする。最後に GPT-4 を使って、サマリーを生成した。

表 4.10: 実験 1 先頭 10 分間の注目トピック抽出結果とサマリー

	post id:1134	post id:1137	post id:1138
注目トピック	topic0:0.9915	topic1:0.9631	topic0:0.9964
topic0 の単語	考える, 貧困, 思う, できる, 必要, 飲む, 定義, 雇用, 資源, 解決		
topic0 のサマリー	貧困問題の解決には資源と雇用の確保が必要		
topic1 の単語	貧困, 思う, 問題, 世界, 制度, 政府, 人口, 足りる, 作る, 膨大		
topic1 のサマリー	貧困は世界的な問題で、政府の制度だけでは解決困難		

表 4.11: 実験 1 中間 10 分間の注目トピック抽出結果とサマリー

	post id:1143	post id:1144	post id:1146
注目トピック	topic0:0.9954	topic0:0.9848	topic0:0.9923
topic0 の単語	十分, 貧困, 考える, 対価, わかる, できる, 定義, 必要, 得る, 住所		
topic0 のサマリー	貧困の定義と解決には十分な対価と住所の確保が必要		

表 4.12: 実験 1 最後 10 分間の注目トピック抽出結果とサマリー

	post id:1150	post id:1151	post id:1152
注目トピック	topic1:0.9496	topic0:0.9615	topic3:0.9635
topic1 の単語	思う, 貧困, 経済的, 出来る, 機会, 困窮, 農村, 好き, 個人的, 種		
topic1 のサマリー	貧困の解決には経済的困窮だけでなく、生活環境も考慮する必要がある		
topic0 の単語	貧困, 十分, 対価, 資源, 考える, 必要, 得る, 豊か, 思う, 感じる		
topic0 のサマリー	貧困問題の解決には十分な対価と豊かな資源が必要		
topic3 の単語	思う, 可能性, 星, 資源, 開発, 新しい, 技術, 発展, 解決策, 実現		
topic3 のサマリー	新しい資源の開発や技術の発展が貧困問題の解決策となる可能性		

表 4.13: 実験 2 先頭 10 分間の注目トピック抽出結果とサマリー

	post id:1160	post id:1168	post id:1162
注目トピック	topic3:0.9194	topic1:0.9599	topic3:0.9330
topic3 の単語	思う, 貧困, 経済的, 出来る, 機会, 困窮, 農村, 好き, 個人的, 種		
topic3 のサマリー	資源配給とアクセスの改善による食料問題の解決		
topic1 の単語	問題, 食品, 飢餓, 日本, 解決, 無料, 現在, 発達, たま, ネット		
topic1 のサマリー	飢餓問題の解決と食品の無料提供についての議論		

表 4.14: 実験 2 中間 10 分間の注目トピック抽出結果とサマリー

	post id:1178	post id:1185	post id:1175
注目トピック	topic1:0.9647	topic2:0.6606	topic2:0.9416
topic1 の単語	思う, 作る, できる, スキル, 先進国, 問題, 良い, 食料, 資源, 個人的		
topic1 のサマリー	先進国によるスキル提供と自立支援の必要性		
topic2 の単語	問題, 解決, 飢餓, 思う, できる, しれる, 先進国, 考える, 多い, 技術		
topic2 のサマリー	飢餓問題の解決と技術的な対策の議論		

表 4.15: 実験 2 最後 10 分間の注目トピック抽出結果とサマリー

	post id:1192	post id:1194	post id:1197
注目トピック	topic2:0.9651	topic2:0.9651	topic2:0.9888
topic2 の単語	思う, の, いい, 先進国, 後進国, いく, できる, 現地, 期待, 名案		
topic2 のサマリー	先進国と後進国の協力と教育による飢餓問題の解決策の探求		

表 4.16: 実験 3 先頭 10 分間の注目トピック抽出結果とサマリー

	post id:1213	post id:1226	post id:1215
注目トピック	topic5:0.9040	topic6:0.5587	topic5:0.9329
topic5 の単語	女性, 男性, 思う, 生理, 抑圧, 弱い, 先天的, 格差, 強い, 社会的		
topic5 のサマリー	男性と女性の生理的な力の格差と社会的抑圧		
topic6 の単語	日本, トイレ, 傾向, 入れる, 感じる, 信頼, 会社, 一時, 労働力, しまう		
topic6 のサマリー	日本のジェンダー問題と社会的傾向		

表 4.17: 実験 3 中間 10 分間の注目トピック抽出結果とサマリー

	post id:1238	post id:1239	post id:1240
注目トピック	topic2:0.9675	topic2:0.9574	topic2:0.9490
topic2 の単語	平等, 女性, GENDER, 思う, 男性, 難しい, くる, 実現, 定義, 感じる		
topic2 のサマリー	ジェンダー平等の実現とその難しさ		

表 4.18: 実験 3 最後 10 分間の注目トピック抽出結果とサマリー

	post id:1253	post id:1255	post id:1261
注目トピック	topic2:0.8761	topic1:0.8251	topic1:0.9720
topic2 の単語	女性, 思う, 男性, くる, 多い, 平等, 男女, 議論, GENDER, 見る		
topic2 のサマリー	男女間の平等とジェンダー問題の議論		
topic1 の単語	平等, 女性, GENDER, 男性, 思う, 実現, 難しい, 社会, 定義, 仕事		
topic1 のサマリー	ジェンダー平等の実現とその難しさ		

表 4.19: 実験 4 先頭 10 分間の注目トピック抽出結果とサマリー

	post id:1276	post id:1277	post id:1279
注目トピック	topic4:0.9544	topic0:0.9865	topic4:0.9908
topic4 の単語	教育, 高位, 質, 提供, 考える, 費用, 状況, できる, 必要, 授業料		
topic4 のサマリー	高質な教育の提供とその費用問題		
topic0 の単語	成長, 幅, 複雑, 方向, 違う, 趣味, 定義, 考える, 内容, 重視		
topic0 のサマリー	教育の目標と学生の成長方向		

表 4.20: 実験 4 中間 10 分間の注目トピック抽出結果とサマリー

	post id:1286	post id:1291	post id:1284
注目トピック	topic2:0.9738	topic2:0.8186	topic2:0.9599
topic2 の単語	教育, 質, 高位, 受ける, 考える, 側, 途上国, 個性, 提供, 思う		
topic2 のサマリー	教育の質とその受け手の視点		

表 4.21: 実験 4 最後 10 分間の注目トピック抽出結果とサマリー

	post id:1299	post id:1304	post id:1296
注目トピック	topic4:0.9627	topic4:0.9643	topic4:0.8284
topic4 の単語	教育, 質, 高位, 考える, できる, 思う, 届く, 子供, 受ける, 以外		
topic4 のサマリー	教育の質とその普及の問題		

表 4.22: 実験 5 先頭 10 分間の注目トピック抽出結果とサマリー

	post id:1309	post id:1308	post id:1312
注目トピック	topic0:0.9845	topic1:0.9870	topic0:0.9883
topic0 の単語	平等, 所得, 教育, 考える, 医療, インフラ, 社会, 単純, 議論, カール大帝		
topic0 のサマリー	社会インフラの不平等性についての議論 (教育、医療、所得)		
topic1 の単語	平等, 差別, 考え, 格差, 中二, 深刻, ちゃう, 話題, 燃える, 終わる		
topic1 のサマリー	格差と差別についての深刻な考察		

表 4.23: 実験 5 中間 10 分間の注目トピック抽出結果とサマリー

	post id:1318	post id:1317	post id:1316
注目トピック	topic2:0.9741	topic1:0.9699	topic1:0.9488
topic2 の単語	貰える, 単位, 時代, 無い, 発展, 子供, 資産, 未来, 持つ, 実現		
topic2 のサマリー	技術発展と資源豊かな時代に期待する未来の実現		
topic1 の単語	不平等, なくす, 小規模, 教育, 議論, 考える, 技術, 資源, される, 問題		
topic1 のサマリー	教育の不平等を減少させるための小規模な問題解決策の議論		

表 4.24: 実験 5 最後 10 分間の注目トピック抽出結果とサマリー

	post id:1327	post id:1331	post id:1324
注目トピック	topic1:0.6143	topic2:0.8455	topic1:0.7428
topic1 の単語	農村部, 地域, 少ない, インターネット, 環境, 人々, 都市, 行く, いいね, 大学		
topic1 のサマリー	農村部と都市の間の教育とインターネット環境の格差		
topic2 の単語	不平等, 国民, 教育, 減少, 大きい, 良い, 解決, 無くなる, 等しい, 一環		
topic2 のサマリー	教育の不平等を減少させるための国民全体の取り組み		

表 4.25: 実験 6 先頭 10 分間の注目トピック抽出結果とサマリー

	post id:1341	post id:1342	post id:1338
注目トピック	topic0:0.8929	topic6:0.7857	topic6:0.7857
topic0 の単語	日本, ガソリン車, 考える, ん, やすい, 不足, 従来, 地形, 使う, 電気		
topic0 のサマリー	日本の電気不足と従来のがソリン車の利便性		
topic6 の単語	開発, 対策, 思う, 電気, 使う, 話題, ガソリン, エネルギー, 電気自動車, 両方		
topic6 のサマリー	電気自動車とガソリン車の開発と対策		

表 4.26: 実験 6 中間 10 分間の注目トピック抽出結果とサマリー

	post id:1353	post id:1354	post id:1358
注目トピック	topic0:0.9920	topic1:0.9875	topic1:0.9875
topic0 の単語	気候変動, 少ない, 考える, 思う, 日本, 電気, 使う, 高い, 効率, 良い		
topic0 のサマリー	日本の電気高騰と気候変動への対策		
topic1 の単語	日本, シェア, サイクル, 普及, 思う, 策, 変化, 温室効果ガス, 部分, 見る		
topic1 のサマリー	日本におけるシェアサイクルの普及と温室効果ガスの削減		

表 4.27: 実験 6 最後 10 分間の注目トピック抽出結果とサマリー

	post id:1361	post id:1362	post id:1363
注目トピック	topic3:0.9326	topic3:0.7319	topic2:0.9578
topic3 の単語	問題, 難しい, シェア, サイクル, 普及, 効果, 温室効果ガス, 見る, 更新, 技術		
topic3 のサマリー	シェアサイクルの普及と温室効果ガス削減の効果		
topic2 の単語	考える, 代替, しれる, 主流, 気候, いく, 出る, くる, 製造業, プラスチック		
topic2 のサマリー	製造業の環境負荷とプラスチック代替品の問題		

本研究は PageRank に基づくスコアリングにより、各実験セッションで最も注目されたトピックが特定された。また、実験結果の内、貧困問題、ジェンダー平等、教育の質、気候変動など、さまざまなトピックが議論されたことを示している。他には、GPT-4 を使用して生成されたサマリーは、LDA モデル出力のトピック内の単語分布より、各トピックの核心を簡潔に表現している。さらに、トピックの単語とサマリーから、議論されたトピックの関連性と重要性が明らかになる。例えば、貧困問題に関する議論では、資源と雇用の確保が重要であることが強調されている。

4.4.4 アンケート評価

実験では、10 分単位で注目内容の抽出作業を行う。アンケートでは、それ抽出した注目内容を使って、参加者がこの時間内に注目した内容は選択肢に当てはまるかどうかを選択する。選択肢の数は、抽出した注目内容の数により 1 から 3 つの間となる。もし、参加者の注目内容は、どの選択肢にも当てはまらない場合は、その他の選択肢があり、自分で注目内容を入力する。実験は 6 回行って、毎回についての選択肢は 3 つある。図 4.1 はその結果を示したものである。

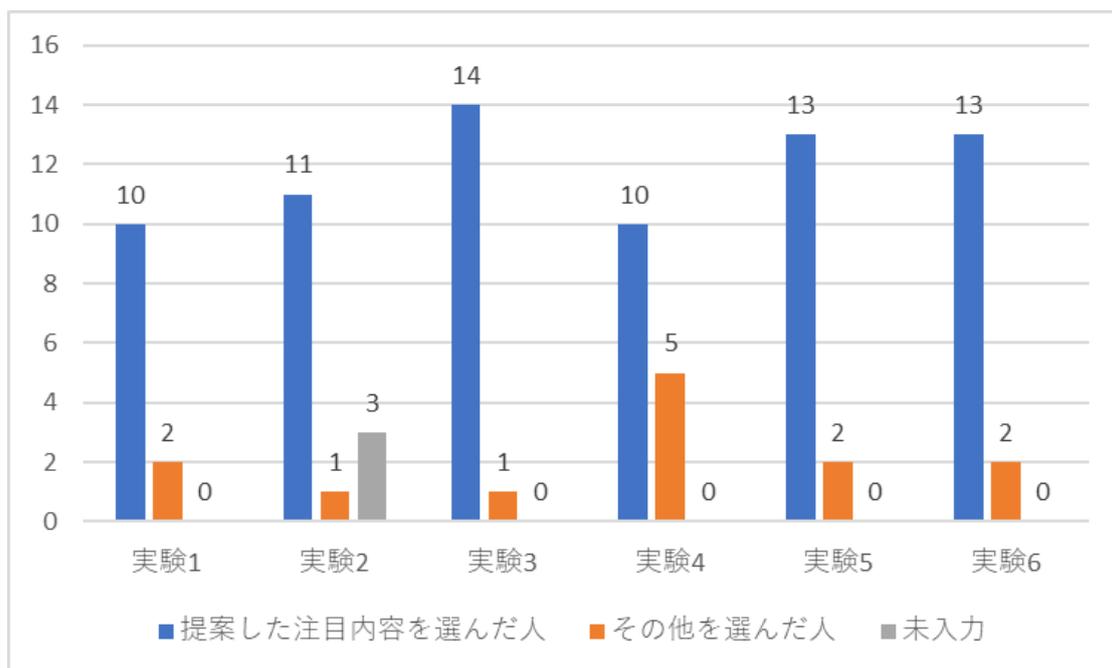


図 4.1: アンケート結果

図 4.1 によると各実験段階で「抽出した注目内容を選んだ人」の数が「その他を選んだ人」と「未入力」の数よりも明らかに多いことがわかる。これは、参加者の多くが与えられた時間内に提供された選択肢の中から彼らが注目した内容を見つけることができたことを示しており、この傾向は全 6 回の実験で一貫している。

しかし、一定の割合の参加者がアンケートの選択肢の中に自分の注目内容を見つけられず、とくに実験 4 では三分の一が「その他」を選択した。今後、これらの参加者が入力した内容を分析し、注目内容の精度向上を検討する。さらに、「未入力」の数が少ないことから、参加者の関与度が高く、ほとんどの人がアンケートで自分の注目内容を示していることがわかる。

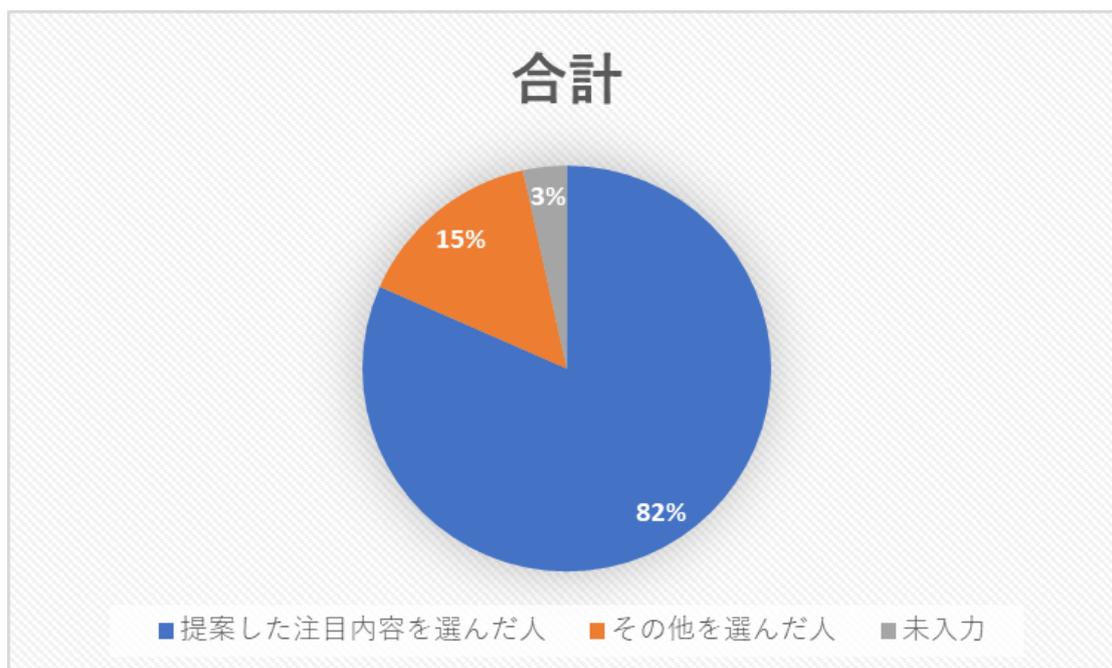


図 4.2: 全体のアンケート結果

本実験では、指定範囲内のポストの PageRank で計算したスコアが Top3 位のポストを注目ポストとしたが、注目度のより上位のポストのトピックで生成されたサマリーがより多く選択されるかを評価するために、図 4.3 を作成した。ここで、PageRank スコアが一番高いポストのトピックで生成されたサマリーを最も注目された内容とし、図 4.3 は最も注目した内容を選択した割合を示したものである。

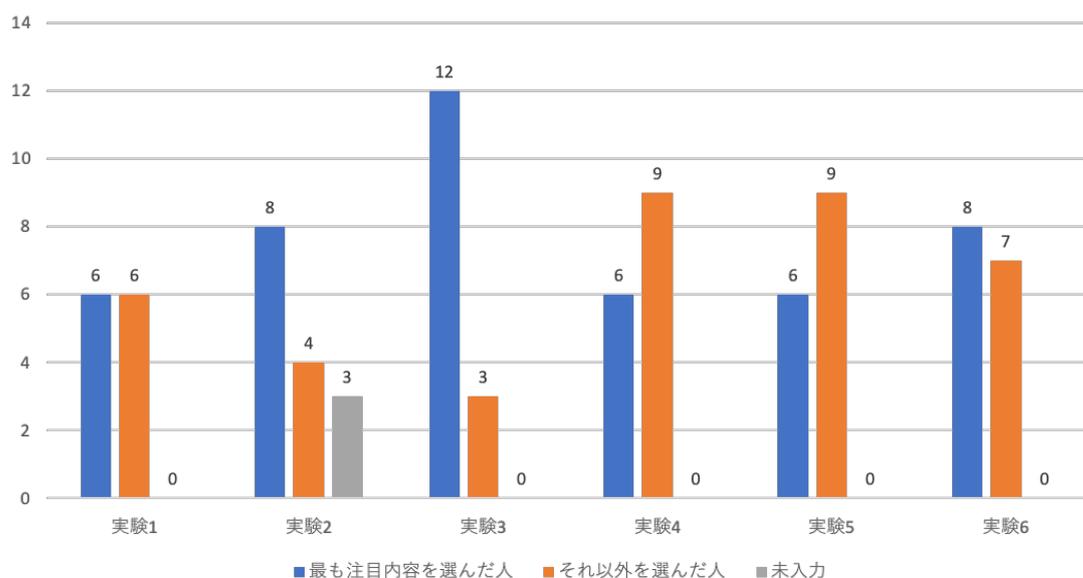


図 4.3: 最も注目した内容のアンケート結果

図 4.3 より、実験 2 と実験 3 では最も注目した内容を選択した人は、他の実験より多かった。その原因として、それぞれのポスト数は 49 と 58 であり、それと比較すると、他の実験のポスト数は 30 程度しかない。このことから、ポストの数は LDA モデルから出力されたトピックの精度に影響する可能性があると考えられる。

4.5 考察

4.5.1 本研究の知見

実験結果によると、PageRank によって特定された注目投稿と LDA モデルによって抽出されたトピックは、参加者のアンケート結果と比較して妥当であった。アンケート結果は、提案手法が参加者の議論中の注目内容を概ね捉えていることを示唆している。ただし、特に実験 4 では、いくつかの参加者が「その他」を選択しており、アンケートが参加者の注目内容を完全にはカバーしていないことが明らかになった。

また、ポスト数が多かった実験 2 と 3 で注目トピックを選択した参加者の割合が高かったことから、投稿数が LDA モデルの精度に影響を与える可能性が示された。これは、オンライン会議の場合はデータ量がトピックモデルの精度に直接影響を与えることを示す重要な知見である。

他にも、GPT-4 による議論内のトピックのサマリー生成は、実験中に参加者が注目したとされるトピックを簡潔にまとめることに成功している。これは、会議中の重要な議論ポイントを効率的に把握するためのツールとしての潜在的な有効性を示している。

よって、提案された手法はオンライン会議における合意形成支援ツールとして有望な結果を示しているが、アンケート項目の改善やデータ量とモデル精度の関係のさらなる検証が今後の課題である。さらに、異なる議論のテーマによってトピックモデルの精度がどのように変わるのか、どのような条件下で最も効果的なサマリーが得られるのかなど、追加研究が必要である。

4.5.2 残された課題

ファシリテーターの行動や判断が会議の結果に大きな影響を与える可能性がある。本研究で開発された AI エージェントは、客観的な参加者間の注目内容を提供し、ファシリテーターの主観的影響を軽減できる可能性がある。しかし、ファシリテーターによる評価で、エージェントが提供する情報や推薦が、実際の会議進行においてどの程度役立つのか、またファシリテーターの負担をどの程度軽減できるのかについては、今後の研究で評価する必要がある。

第5章 まとめと今後の課題

5.1 まとめ

本研究では、オンライン会議の合意形成を対象に、情報処理支援の視点から、会議参加者の注目内容をファシリテーターへ提供する会議支援エージェントを開発し、その有効性を実験的に評価した。

このエージェントは、LDA モデルと PageRank アルゴリズムの組み合わせにより、会議中の重要なポストとトピックを効果的に特定できた。これにより、会議のキーポイントを把握し、参加者の議論を支援することが可能となった。また、抽出されたトピックに基づいて、GPT-4 を用いてサマリーを生成した。これにより、会議の要点を簡潔にまとめ、参加者が議論の全体像を理解しやすくなった。

その後エージェントを評価するために、実際の会議実験を行った。実験では、オンラインプラットフォーム上で会議を行い、会議データを収集した。収集したデータは、LDA モデルと PageRank を用いて分析され、議論内容の抽出が行われた。参加者のアンケート結果と比較し、AI エージェントの効果を評価した。

その結果、AI エージェントによって特定された注目投稿と抽出されたトピックは、参加者のアンケート結果と一致しており、提案手法が参加者の議論中の注目内容を捉えていることが示された。特に、ポスト数が多い会議では LDA モデルの精度が向上する傾向が見られた。

以上のことから、本研究で開発された AI エージェントは、オンライン会議における合意形成を支援する有効なツールであることが示された。参加者からの客観的な注目内容をファシリテーターへ提供することができた。今後の研究では、異なる議論テーマや会議環境でのエージェントの適用性と効果をさらに検証することが期待できる。

5.2 今後の課題

5.2.1 大人数の会議実験

大人数の会議では、参加者の背景、専門知識、意見の多様性が増加する。この多様性は、議論の複雑性を高め、合意形成のプロセスをより複雑にする。AI エー

エージェントがこのような多様な環境でどのように機能するかを理解することは、その汎用性と適応性を評価する必要がある。

5.2.2 エージェント処理速度の向上

最適な LDA モデルのパラメータを探すためには、それぞれのパラメータを試し、一貫性スコアを評価しなければならない。筆者で使われた Macbook Air、CPU:Apple M2、メモリ:8GB の場合は、5分ほど時間がかかり、リアルタイムでの会議の注目内容を評価することが難しい。そのために、高速なパラメータ特定アルゴリズムを適用して、エージェント処理速度を向上させることが必要である。

5.2.3 ポストに対する重みづけ

注目ポストを計算する時、PageRank を使って、計算を行った。その時、ディフォルトで全てのポストを同じ重みで計算をした。しかし、実際の会議では、発言の重みはそれぞれにより異なる。例えば、会議のとき、一般市民の発言と専門家の発言の重みは同じではない。また、時間にも重みは影響する、1時間前の発言と現在の発言の重みは異なる。それぞれの重みをどのように定義し利用するかは、今後の課題である。

5.2.4 ファシリテータでの評価

ファシリテーターの役割は、会議の円滑な進行を確保し、効果的な意思決定や合意形成を促進することにある。そのため、ファシリテーターの行動や判断が会議の結果に大きな影響を与える可能性がある。本研究で開発された AI エージェントは、ファシリテーターを支援し、その主観的影響を軽減することを目的としているが、エージェントの効果を正確に評価するためには、ファシリテーター自身の評価が不可欠である。

ファシリテーターによる評価は、エージェントが提供する情報や推薦が、実際の会議進行においてどの程度役立つか、またファシリテーターの負担をどの程度軽減できるかを判断する上で重要な指標となる。この評価を通じて、エージェントの機能改善点や、ファシリテーターとエージェントとの相互作用の最適化方法を特定することができる。したがって、ファシリテーターによる評価は、エージェントの有効性を検証し、オンライン会議における合意形成をより効果的に支援するために必要である。

研究業績

- 王 治中, 谷 文, 太田 光一, 長谷川 忍. “オンライン会議における合意形成のためのファシリテーター支援エージェント”, 2023年度 第4回 合意と共創 (Consensus) 研究会, 北海道大学. (2024 in press)

謝辞

本研究の遂行にあたり、指導教員である北陸先端科学技術大学院大学の長谷川研究室、長谷川 忍教授と谷 文助教から多大なご指導と助言をいただき、この度、深謝致します。

中間審査会では金子教授、白井准教授から、本論文の作成にあたり、副査として適切なお助言を賜りました。ここに深謝の意を表します。

最後に、長谷川研究室の皆様には、本研究の遂行にあたり多大なお助言、ご協力頂きました。ここに感謝の意を表します。

参考文献

- [1] Fangfang HU. “The Process of Consensus Building seen in Group Discussions by Students Learning Japanese as a Second Language: Based on Oral Materials during Peer-Reading Studies.” 一橋日本語教育研究, vol. 4 (Mar. 2016), pp. 127–136. URL: <https://cir.nii.ac.jp/crid/1050568617183515392>.
- [2] Terri L. Griffith, Mark A. Fuller, and Gregory B. Northcraft. “Facilitator Influence in Group Support Systems: Intended and Unintended Effects.” *Inf. Syst. Res.* vol. 9 (1998), pp. 20–36. URL: <https://api.semanticscholar.org/CorpusID:39682852>.
- [3] Hikaru Ishizuka, Shun Shiramatsu, and Keiko Ono. “Prototyping Agents for Resolving Opinion Biases Toward Facilitating Sublation of Conflict in Web-based Discussions.” *2022 IEEE International Conference on Agents (ICA)*. 2022, pp. 18–23. DOI: [10.1109/ICA55837.2022.00010](https://doi.org/10.1109/ICA55837.2022.00010).
- [4] 国土交通省. “新型コロナウイルス感染症の影響を受けた国際会議に係る調査等業務報告書.” Mar. 2023. URL: <https://www.mlit.go.jp/kankocho/content/001580267.pdf>.
- [5] 孝行 伊藤, 桂英 藤田, 徳朗 松尾, and 直樹 福田. “エージェント技術に基づく大規模合意形成支援システムの創成：自動ファシリテーションエージェントの実現に向けて.” *Journal of the Japanese Society for Artificial Intelligence*, vol. 32, no. 5 (Sept. 2017), pp. 739–746. ISSN: 21882266. DOI: [10.11517/jjsai.32.5_739](https://doi.org/10.11517/jjsai.32.5_739). URL: <https://cir.nii.ac.jp/crid/1390567172581532288>.
- [6] 土木学会誌編集委員会 編. “合意形成論：総論賛成・各論反対のジレンマ.” 土木学会誌叢書 ; 2. 土木学会, 2004.3. URL: <https://ndlsearch.ndl.go.jp/books/R100000002-I000007353894>.
- [7] 総務省. “新型コロナウイルス感染症が社会にもたらす影響.” 令和2年版情報通信白書(PDF版). 2020. URL: <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r02/pdf/n2300000.pdf>.
- [8] 孝治 長谷川, コウジ ハセガワ, 瑠理奈 岩瀬, ルリナ イワセ, Koji Hasegawa, and Rurina Iwase. “対面, テキストベース, ビデオチャットでのコミュニケーションがソーシャル・サポートと well-being に及ぼす効果：新型コロナウイルス感染拡大状況における検討.” 駒澤大学心理学論集：KARP,

- vol. 24 (Mar. 2022), pp. 25–39. URL: <https://cir.nii.ac.jp/crid/1050291768485575168>.
- [9] Yuki ANZAI and Shoko AOKI. “Workshop Practitioners’ Recognition of Difficulties in Facilitation.” *Japan Journal of Educational Technology*, vol. 42. no. 3 (Jan. 2019), pp. 231–242. ISSN: 1349-8290. DOI: [10.15077/jjet.42073](https://doi.org/10.15077/jjet.42073). URL: <https://cir.nii.ac.jp/crid/1390845713050130432>.
- [10] 浦山昌志. “令和版！進化した4つのファシリテーションスキル.” 日本フューチャーラーナーズ協会, (2021). URL: <https://japan-future-learners.or.jp/columns/urayamainterview/>.
- [11] Takayuki Ito, Shota Suzuki, Naoko Yamaguchi, Tomohiro Nishida, Kentaro Hiraishi, and Kai Yoshino. “D-Agree: Crowd Discussion Support System Based on Automated Facilitation Agent.” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34 (Apr. 2020), pp. 13614–13615. DOI: [10.1609/aaai.v34i09.7094](https://doi.org/10.1609/aaai.v34i09.7094).
- [12] Ryosuke Kinoshita and Shun Shiramatsu. “Agent for Recommending Information Relevant to Web-based Discussion by Generating Query Terms using GPT-3.” *2022 IEEE International Conference on Agents (ICA)*. 2022, pp. 24–29. DOI: [10.1109/ICA55837.2022.00011](https://doi.org/10.1109/ICA55837.2022.00011).
- [13] Yuki Yoshimura, Shun Shiramatsu, and Takeshi Mizumoto. “Semi-automatic Summarization of Spoken Discourse for Recording Ideas using GPT-3.” *IIAI Letters on Informatics and Interdisciplinary Research*, vol. 3 (Feb. 2023), p. 1. DOI: [10.52731/liir.v003.070](https://doi.org/10.52731/liir.v003.070).
- [14] Takayuki Ito. “Towards Agent-based Large-scale Decision Support System: The Effect of Facilitator.” Jan. 2018. DOI: [10.24251/HICSS.2018.047](https://doi.org/10.24251/HICSS.2018.047).
- [15] 奥原 俊 and 伊藤 孝行. “ビデオ会議におけるエージェント介入による意見の変化の分析.” *人工知能学会全国大会論文集*, vol. JSAI2023 (2023), 2T5GS504–2T5GS504. DOI: [10.11517/pjsai.JSAI2023.0_2T5GS504](https://doi.org/10.11517/pjsai.JSAI2023.0_2T5GS504).
- [16] Chunsheng Yang, Wen Gu, Takayuki Ito, and Xiaohui Yang. “Machine learning-based consensus decision-making support for crowd-scale deliberation.” *Applied Intelligence*, vol. 51 (July 2021). DOI: [10.1007/s10489-020-02118-z](https://doi.org/10.1007/s10489-020-02118-z).
- [17] Katsuhide Fujita, Takayuki Ito, and Mark Klein. “Enabling Large Scale Deliberation Using Ideation and Negotiation-Support Agents.” June 2017, pp. 360–363. DOI: [10.1109/ICDCSW.2017.75](https://doi.org/10.1109/ICDCSW.2017.75).
- [18] BLEI D. M. “Latent dirichlet allocation.” *The Journal of Machine Learning Research*, vol. 3 (2003), pp. 993–1022. URL: <https://cir.nii.ac.jp/crid/1571698601169800320>.

- [19] Rickjin (靳志). “LDA 数学八卦.” Feb. 2013. URL: <https://github.com/zpfbuaa/Paper-1/blob/master/%E4%B8%BB%E9%A2%98%E6%A8%A1%E5%9E%8B/LDA%E6%95%B0%E5%AD%A6%E5%85%AB%E5%8D%A6/LDA%E6%95%B0%E5%AD%A6%E5%85%AB%E5%8D%A6.pdf>.
- [20] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. “The PageRank Citation Ranking : Bringing Order to the Web.” *The Web Conference*. 1999. URL: <https://api.semanticscholar.org/CorpusID:1508503>.
- [21] Despina Christou. “Feature extraction using Latent Dirichlet Allocation and Neural Networks: A case study on movie synopses.” (Apr. 2016).
- [22] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. “Attention Is All You Need.” 2023. arXiv: [1706.03762](https://arxiv.org/abs/1706.03762) [cs.CL].
- [23] 初 CV. “Transformer 模型解 (解最完整版).” Dec. 2020. URL: <https://www.overleaf.com/project/64ae0034e3a0b6a5583205ac>.
- [24] OpenAI et al. “GPT-4 Technical Report.” 2023. arXiv: [2303.08774](https://arxiv.org/abs/2303.08774) [cs.CL].
- [25] Enes Zvornicanin. “When Coherence Score Is Good or Bad in Topic Modeling?” Baeldung. 2023. URL: <https://www.baeldung.com/cs/topic-modeling-coherence-score>.