

LLMを用いた自己改良型プロンプトによる 組織内の認知共有支援システムの開発

Development of an Organizational Cognitive Sharing System with Self-Refining LLM Prompts

伊藤誠二^{1*} 長澤史記¹ 白松俊¹
Seiji Ito¹ Huminori Nagasawa¹ Shun Shiramatsu¹

¹ 名古屋工業大学

¹ Nagoya Institute of Technology

Abstract: In organizations, mutual understanding of members' work is essential for smooth collaboration, yet the sharing of progress and issues often depends on individuals' voluntary reporting. This study developed a system that supports the formation of transactive memory (TM) by using a chatbot powered by a large language model (LLM) to collect progress updates and share them as reports. A question regeneration method was also proposed, evaluating prompts from perspectives such as "contribution to TM," "psychological load," and "self-efficacy." Experimental results showed that the effect of the regeneration method was limited, but continued use of the system led to a statistically significant improvement in understanding of projects and domains ($p = 0.0312$). These findings demonstrate that the system itself serves as an effective tool for promoting knowledge sharing and mutual understanding within organizations.

1 はじめに

1.1 本研究の背景

現代の組織においては、メンバー間での業務内容や保有知識の相互理解が効率的な協働と成果創出に不可欠である。しかし、こうした進捗や課題の共有は、多くの場合、個々のメンバーの自発的な発信に依存しており、情報共有の頻度や質にばらつきが生じ、組織内での認識の偏りや、重要情報の伝達遅延といった問題を引き起こしている。このような課題に対応するため、本研究ではトランザクティブ・メモリ (Transactive Memory, TM) の概念に着目した。TM とは、Wegner ら [1] によって提唱された「誰が何を知っているか」についての共有認識を指し、チーム内での知識分散と統合を効率化する機能を持つ。本研究ではこの概念を「誰が今何をしているのか」「タスクやプロジェクトの現状はどうなっているのか」といった活動状況の共有へと拡張し、その認知の支援を試みる。

こうした共有認識がチームや組織レベルで機能する際には、しばしばトランザクティブ・メモリ・システム (Transactive Memory System, TMS) という枠組

みで捉えられる。TMS とは、個人が持つ知識や活動に関する情報が、組織内で誰にどのように分散され、どのように活用されるかを含んだ認知的かつ社会的なシステムであり、円滑な協働を支える重要な基盤となると考える。

そこで本研究では、TM の共有認識を形成・活用可能にするため、TMS を構成する要素の一部を実装した情報共有システムを開発した。具体的には、大規模言語モデル (LLM) を活用したチャットボットが Slack 上で個人の進捗を聞き取り、レポート形式で共有することで、メンバー間の認知を支援する仕組みを構築した。

さらに、本システムによる TMS 形成においては、この「聞き出し」プロセスの質が極めて重要である。共有される情報の内容は、いかに有用な発言を引き出せるかに依存しており、そのためには、TMS 形成に資する内容を含む質問を提示すると同時に、ユーザが心理的負荷なく応答できるような自然な対話設計が求められる。しかし、単に LLM に対してプロンプトを与えるだけでは、これらの要件を満たす質問が安定的に生成されるとは限らない [2]。そこで本研究では、このシステムのための進捗聞き取りに用いる質問文の生成手法に焦点を当てた。ユーザがより自然に情報を提供できるよう、提案手法では「TM 形成への寄与」「心理的負荷」「文脈の自然さ」「自己効力感」の 4 観点から質

*連絡先: 名古屋工業大学大学院工学研究科
〒466-0061 愛知県名古屋市昭和区御器所町
E-mail: s.ito.303@stn.nitech.ac.jp

問を評価し、再生成を行う仕組みを導入した。

1.2 本研究の目的

本研究の主たる目的は、開発したシステムを実運用環境に適用し、生成された質問およびレポートがチーム内の TM 形成や相互理解の促進に与える効果を検証することである。

加えて、効果的かつユーザ満足度が高い情報収集の実現を目的として、質問生成手法の開発にも焦点を当てた。具体的には、(1) ユーザの心理的負担を軽減し、(2) 自然な対話の流れの中で回答意欲を高める質問を生成しながら、(3) 組織内の「誰が何をしているか」という TM 形成に有用な情報を効率的に収集できる手法の確立を目指した。これによりシステムとユーザ間の対話の質を向上させつつ、組織内の情報共有と相互理解を促進するための技術の開発を試みた。

2 関連研究

2.1 トランザクティブメモリシステム (TMS)

大沼 [3] は、企業、組織内において TMS が部門単位での成果にどのような影響を与えるかを定量的に検証した。結果として、部門内の対面コミュニケーションが部門内 TMS にプラスの影響を与えるということを明らかにした。

宇佐美 [4] は学校組織における TMS 構築の実践的研究として、小学校教職員を対象とした研修プログラムの開発とその効果検証を行った。同研究では、教職員同士が「誰が何を知っているか」を共有し、学年を超えたつながりを深めることで組織内の協働的関係の構築を目指した。具体的には、Lewis ら [5] の TMS 測定尺度における 3 つの下位次元（専門性・調整・信頼性）を基盤として、「コネクト★カード」といったツールを活用した研修を 4 回実施した。その結果、特に「専門化」と「調整」の観点で統計的に有意な向上が確認され、教職員間の情報共有のスムーズさや相互理解の促進が観察された。

2.2 対話システム・チャットボットについて

角森ら [6] は、ユーザに長く使用される対話システムを開発するために、ユーザ情報を自然に活用することを試みた。具体的には、任意のユーザ情報と対話文脈を踏まえたシステム発話を含む SUI コーパスを構築し、これを用いて事前学習済み発話生成モデルを Fine-tuning するとともに、大規模言語モデル (LLM) に In-Context

Learning を適用した。その結果、現在のトピックに関係なくユーザ情報を取り込んだ自然な発話生成が可能になり、ユーザ満足度が向上した。

Madaan ら [2] は、LLM が自身の生成した発話に対してフィードバックを与え、それに基づいて反復的に改善する SELF-REFINE を提案した。同手法では、関連性、一貫性など 10 個の観点から LLM が自身の出力を評価し、そのフィードバックをもとに発話を再生成することで、ユーザにとってより良い対話を実現した。

2.3 本研究の独自性・立ち位置

先行研究では、組織内の TMS の重要性やそれに影響を与える要因について定量的な分析が行われている一方で、TM 及び TMS を積極的に形成・構築するためのインタラクティブな支援手法に関する研究は限定的である。また、対話システム・チャットボットに関する研究では、TM の形成を目的とした対話設計は十分に検討されていない。また対話システムの発話の評価・再生成する取り組みはあるが、TM 形成のための聞き取り質問に限定してこれを行う研究はない。

本研究は、チャットボットによる進捗聞き取りとその情報の可視化を通じて、TM の形成を支援する情報共有システムを提案するものである。さらに、ユーザの心理的・認知的負荷や質問の文脈整合性を考慮した質問再生成手法を導入し、対話システムとしての質向上と TM 形成支援の両立を図る点に独自性がある。

3 提案手法・システム概要

組織において、メンバー間の活動状況や知識の相互理解が、連携と円滑化に必要であると考えられるが、その共有は個人の自発的な発信に依存する面があり、情報の伝達遅延といった課題を引き起こしている。本研究では、このような状況を改善し、TM の形成を支援することを目的として、非同期コミュニケーションツールである Slack を用いた情報共有システムを開発した。

本システムは、Slack 上で稼働するチャットボットがユーザに対して定期的に進捗や課題を質問し、その回答を他のアクティビティと統合してレポートを生成・共有することで、チーム内での認知の共通化を図るものである。ここでいうアクティビティとは、組織全体のチャンネルおよび各プロジェクトチャンネルにおける投稿を指し、これらの情報と個人の進捗聞き取り結果を統合することで、包括的な組織活動の把握を可能にしている。

このような聞き取りを通じた情報共有は、TMS 形成において有効であると考えられるが、その効果を十分

に引き出すためには、ユーザから有用な情報を自然に引き出す対話設計が不可欠である。しかし、単に LLM に質問生成を委ねただけでは、ユーザの心理的負担を考慮しつつ TM に資する情報を安定的に得ることが難しい。そこで本研究では、4つの観点から、質問文を自動的に評価・再生成する手法を提案する。

3.1 質問文の自己改良手法

3.1.1 手法の概要

本研究では、ユーザから TM 形成のための情報を収集しつつ、回答に対する心理的負担を低くする、答えなくなる質問を生成する手法を提案する。システムは図 3 のプロンプトにより生成された質問に対して定義した 4 つの評価観点に基づいてスコアリングを行い、その条件を満たさない場合には再生成を行う。以下に示す 4 つの観点から 5 点満点で評価を行い、そのスコアが指定された条件を満たすまで再生成を繰り返す。質問文がスコア条件を満たす、あるいは再生成回数の上限に達した時点で確定される。また、スコアリングには Claude-3.7-sonnet を使用する。

3.1.2 質問の評価基準

質問文の評価には以下の 4 項目を用いる：

1. **TM 形成への寄与**：質問に対する回答が、他のチームメンバーにとってプロジェクトの進捗や組織の状況理解に資する内容となる可能性が高いか。重複情報の提示や、細かすぎる技術的説明の誘導となっていないか評価する。
2. **心理的・認知的負荷**：質問が単一の問いで構成されており、複雑な条件を含まず、自然に回答しやすい構造となっているかを評価する。
3. **自己効力感の向上**：質問文が共感的な表現を含み、ユーザの自己効力感を高めるような内容となっているかを評価する。
4. **文脈の整合性**：文脈の自然さや、ユーザが関与していないプロジェクトについて断定を含んでいないかなどを評価する。

これらの項目は Madaan ら [2] の再生成における評価基準を参考にして作成した。このような観点からシステム発話を評価・再生成することで、よりユーザが答えやすい、かつ組織にとって役に立つ情報が得られる質問が生成できると考えた。具体的なプロンプトは図 1 のようになり、Slack の履歴、対話履歴、各項目のスコアリングの基準が含まれている。

```
# 指示
答えやすく組織にとって意味があるか評価させる
# 評価観点 (各 5 点満点で 0.1 点区切り)
1. TM 的評価, スコア別評価基準
2. 質問の明確さと単純さ, スコア別評価基準
3. 自己効力感の向上, スコア別評価基準
4. 文脈整合性, スコア別評価基準

# 評価対象の質問:
# これまでの対話履歴:
# ユーザが参加しているプロジェクトのアクティビティ
# 出力フォーマット:
- TM 的評価 5 点満点 (int) とその評価理由 (str)
- 心理的・認知的負荷 5 点満点 (int) とその評価理由 (str)
- 自己効力感の向上 5 点満点 (int) とその評価理由 (str)
- 文脈整合性 5 点満点 (int) とその評価理由 (str)
## 出力例 (json)
```

図 1: 質問評価のためのプロンプト

また、再生成の終了タイミングにおいては、(1) 各項目におけるスコアがすべて 3.5 点以下にならないこと (2) 4 項目の平均スコアが 4.0 点以上であること という 2 点の条件を全て満たす場合とする。これにより、著しく質の低い質問を排除しつつ、全体的に質の高い質問のみがユーザに提示されることを担保する。

3.2 システム概要

3.2.1 全体構成と使用技術

本システムは、Slack 上で稼働するチャットボットとして、ユーザに「最近のタスク」や「直面中の課題」などを定期的に質問する。ユーザは Slack の日報チャンネルでこれに回答し、その内容は他のアクティビティと統合され、翌日に要約レポートとしてチーム全体に共有される。システムの構成図は図 2 のようになっている。

レポートには個人の進捗や関与プロジェクト、各チャンネルの活動概要が含まれ、チーム内で「誰が何をしているか」の把握を支援する。このような情報共有を通じて、TM の形成を促進することを目的としている。

本システムは、AWS Lambda 上に構築されており、Slack 上のアクティビティは DynamoDB に、対話履歴や発話生成過程のログは S3 にそれぞれ保存される。対話の生成およびレポートの作成には大規模言語モデル (LLM) のうち、Anthropic 社の Claude-3.7-sonnet を使用した。システムの構成図は図 2 のようになっている。

3.2.2 対話モジュール

対話モジュールは、ユーザに対してパーソナライズされた質問を生成し、Slack 上で自然なコミュニケー

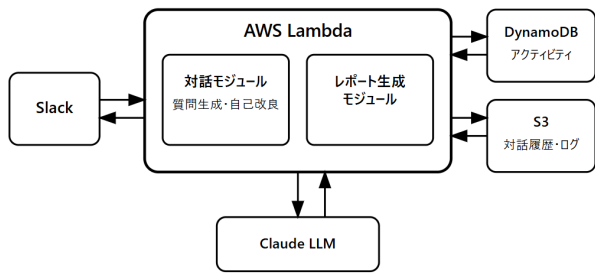


図 2: システム構成図

ションを促す役割を担う。質問は DynamoDB に蓄積された直近の Slack アクティビティから、ユーザと関連性の高いトピックを抽出した上で生成される。質問生成の際には、ユーザへの心理的・認知的負荷を軽減し、自己効力感を高めることを意図したプロンプト設計が行われている。また、本システムでは Slack の各チャンネルが各プロジェクトに対応すると仮定して聞き取りプロンプトを設計した。図 3 に使用したプロンプトの構成を示す。

```

# 指示
あなたは、TM 構築のため、チームの状況を聞き取るための「質問文」を生成する AI です。
(省略)

# 質問に含める要素
## パーソナライゼーション
## トランザクティブメモリの構築
## ラポール形成

# 質問文の条件
- 質問文は必ず 1 つにすること。? (クエスチョンマーク) の数を 1 つにすること
(省略, 2 回目以降の質問に関して深掘りや話題の変更に
関する条件が記述してある)

## 質問例
### ユーザの関わっているプロジェクト、作業が明確な場合
### ユーザの関わっているプロジェクト、作業を推測できない場合

# slack 上のアクティビティ
- (各チャンネルの投稿履歴などの要約) # これまでの対話履歴

```

図 3: 質問生成のためのプロンプト

生成された質問は、ユーザごとの日報チャンネルに対して定期的に投稿される。対話例を図 4 で示す。

対話の終了タイミングとしては、2 つあり、(1) ユーザが終了を希望した場合 (2) システムがこの対話でもう聞くことはない判断した場合 とした。以上のいずれかの条件を満たした場合、対話が終了する。ま

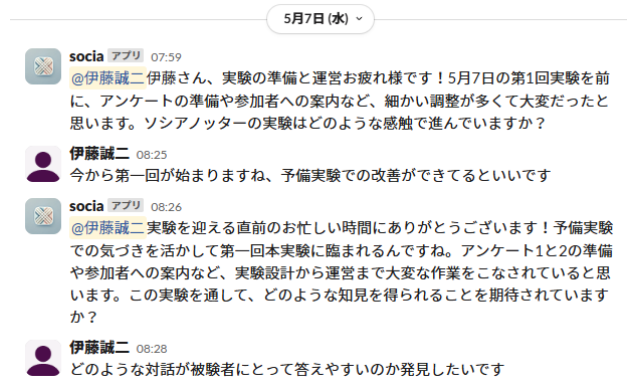


図 4: Slack 上の対話例

た、これらの判定も LLM によって行われる。

3.2.3 レポート生成と配信

ユーザへの聞き取りから 24 時間後に、個別および全体の Slack アクティビティをもとに、その期間における全体の活動を要約したレポートが生成され、メンバー全体に配信される。

レポートは各チャンネルにおけるプロジェクトの動向、および個人ごとの進捗や関与プロジェクトに関する要約で構成される。これにより、メンバーは自身が関与していない領域の活動についても把握でき、TM 形成の促進が期待される。

4 実験

4.1 実験目的と仮説

本実験では、以下の 2 つの観点から提案システム及び発話生成手法の有効性を検証することを目的とする。

仮説 1 TM 形成、心理的・認知的負荷、自己効力感、文脈整合性の 4 つの観点から質問文を評価・再生成することにより、ユーザにとってより良好な対話体験を提供できる (心理的負担の軽減、対話への積極的参加の促進、TM 形成に資する情報の効果的な収集)

仮説 2 システムの継続的な利用により、組織内における TM 形成が促進される。

4.2 実験設計

本実験は、筆者が所属する研究室に併設されたベンチャー組織 9 名から筆者を除いた 8 名のうち、実験を

完遂した7名のデータを対象として評価を行った。実験では、開発した情報共有システムを用いて、定期的な対話による進捗聞き取り（質問の投稿）と、その内容を要約したレポートの配信を行った。質問は各参加者の Slack 上の日報チャンネルに投稿され、回答内容と Slack アクティビティをもとに翌日にレポートが生成・配信されるサイクルを、計4回にわたり実施した。このシステムは以下のスケジュールで動作した：

- 質問投稿日 (Slack で進捗を聞き取る日)：5月7日、12日、16日、21日
- レポート配信日 (前日の聞き取り内容とアクティビティを要約・共有する日)：5月8日、13日、17日、22日

質問は各メンバーの日報チャンネルに投稿され、それに対する回答が対話として記録された。各回の聞き取り終了後、期間内の活動を要約したレポートが全体に共有された。

本実験では、ベースライン手法（3.2節で述べた4観点による評価・再生成を行わない従来の質問生成）と提案手法（4観点による評価・再生成を行う質問生成）の2つのアプローチを比較検証した。7名の被験者を4人と3人（それぞれグループA、グループBとする）に分けて実験を行った。グループAは前半2回の実験でベースライン手法、後半2回で提案手法を用いたシステムと対話を行った。また、グループBはグループAと期間を入れ替えて実験を行った。

4.2.1 アンケートの概要

システムの有効性を多角的に評価するため、以下の2種類のアンケートを実施した。いずれもリッカート尺度（7件法）を用いて評価を行った。

アンケート1：TMの形成度合いおよびレポートの有用性について評価するものであり、LewisらのTMS研究[5]を参考にしつつ、本研究の文脈に適合するよう一部改変した。

アンケート2：対話システムに関する印象やユーザビリティに焦点を当てたものであり、角森らの研究および東中ら[7]による対話システムのコンペティションの評価基準などを基に作成した。

アンケート1は、実験開始前、第2回と第3回の間（実験中）、および実験終了後の計3回実施した。また、実験中と終了後では、TMに関する項目に加えてレポートに対する評価項目も含めて回答を求めた。

アンケート2は対話終了直後に実施し、全4回の対話それぞれに対して実施された。これにより、質問の

質や回答時の印象、心理的負担に関する定量的評価を取得した。以下の表1及び表2にアンケートの項目を示す。

また、本研究では、これらのアンケート結果に対してウィルコクソンの符号付順位検定を用いて統計的分析を行った。これはサンプル数が少なく、データの正規性が仮定できない条件下でも適切に差の有無を検出できるためである。有意水準は原則として5% ($p < 0.05$) とし、必要に応じて10% ($p < 0.10$) で有意傾向も補足的に検討した。

表1: アンケート1：TM形成およびレポートに関する評価

設問	質問項目
1	組織内でどのメンバーがどのプロジェクトや分野に携わっているかを把握している。
2	組織ではメンバーがお互いの活動を理解した上で連携していると感じる。
3	メンバーの誰がどのような専門知識・ノウハウを持っているか把握している。
4	自分が組織内で何をすべきか理解している。
5	システムから送られたレポートは、自分自身の活動内容を正確に反映していたと感じる。
6	レポートは、他のメンバーの活動内容も正確に反映していたと感じる（自分が把握している範囲において）。
7	レポートは、組織の全体的な活動や動向を理解するのに役立ったと感じる。

表2: アンケート2：対話システムに関する評価

設問	質問項目
1	システムの発話は自然であったか。
2	システムの発言は親しみやすいものだったか。
3	このシステムとまた話したいか。
4	システムからの質問は現在の自分の状況に当てはまるものだったか。
5	システムからの質問は分かりやすく、答えやすいと感じたか。
6	システムからの質問に答えるのは負担にならなかったか。

4.3 実験結果

4.3.1 対話システムの評価（アンケート2）

仮説1の検証として、対話システムに関する評価を行った。提案手法とベースラインを比較した結果を表

3に示す。表はそれぞれの設問の平均値を示している。設問1～4（システムの自然性、親しみやすさ、再利用意向、文脈適合性）については提案手法の方が高い数値を示した一方、設問5～6（回答のしやすさ、心理的負担の軽減）についてはベースラインの方が高い数値を示した。

表 3: 対話システムの評価結果（アンケート 2）

設問	提案手法	ベースライン	p 値
1	5.36	5.21	0.813
2	5.50	5.21	0.656
3	5.14	4.64	0.500
4	5.43	5.29	0.813
5	5.43	5.64	0.500
6	5.36	5.50	1.000

また、ウィルコクソンの符号付順位検定では、いずれの設問においても統計的に有意な差は認められなかった（ $p > 0.05$ ）。

加えて、仮説1の「TM 形成に資する情報の効果的な収集ができたか」という部分の評価を行った。提案手法の再生成プロセスを経た24件の質問への回答と、「提案手法だが再生成していない質問への回答」および「ベースライン質問への回答」100件の計124件を対象に補完的な定性分析を実施した。各回答を「組織にとって有用な具体的情報を含むか」の観点から筆者が評価した結果、再生成プロセスを経た回答では有用情報を含む割合が50%（12/24件）、それ以外では56%（56/100件）であった。

4.3.2 TM 形成の評価（アンケート 1）

TM の形成に関する評価結果を表4に示す。これも7人の回答の平均値を示している。実験前、実験中、実験後の段階的な数値の向上が確認され、特に実験前後の比較において顕著な改善が見られた。

表 4: TM 形成の評価結果（アンケート 1）

設問	実験前	実験中	実験後	p 値（前→後）
1	2.57	4.29	5.57	0.0312*
2	2.14	3.29	5.00	0.0625 †
3	2.71	4.14	5.43	0.1250
4	3.14	4.14	5.43	0.0625 †

(* $p < 0.05$, † $p < 0.10$)

各設問の平均値は、設問1（プロジェクト・分野の把握）では2.57から5.57へ、設問2（メンバー間の連携

理解）では2.14から5.00へ、設問3（専門知識・ノウハウの把握）では2.71から5.43へ、設問4（自身の役割理解）では3.14から5.43へと、すべての項目において大幅な向上を示した。

実験前後の比較においてウィルコクソンの符号付順位検定を実施した結果、設問1では統計的に有意な向上が認められ（ $p = 0.0312$ ）、設問2および設問4では向上する傾向が確認された（それぞれ $p = 0.0625$ ）。設問3についても数値的な向上は見られたが、統計的有意差には至らなかった（ $p = 0.1250$ ）。

また、設問5～7のレポートの評価項目についてはアンケート2回分の平均をまとめて集計したところ、すべての設問で5.5となった。

なお、提案手法期間とベースライン期間のTM 向上幅を比較した追加分析についても実施した（詳細は考察にて述べる）。結果を表5に示す。灰色部分が提案手法を使用した期間内での向上幅を表している。全ての設問において「提案手法を用いた期間の向上幅 \geq ベースライン期間の向上幅」という結果が得られた。また、同グループのそれぞれの手法を用いた期間においてウィルコクソンの符号付順位検定を行ったが、有意差は確認できなかった。

表 5: アンケート 1 における平均値の向上幅
（期間・設問別、グループ比較）

グループ	設問	前 → 中	中 → 後
A	1	2.00	2.00
	2	1.25	3.00
	3	1.75	2.00
	4	1.25	2.25
B	1	1.33	0.33
	2	1.00	0.00
	3	1.00	0.33
	4	0.67	0.00

4.4 仮説検証

本節では、4.1節で提示した仮説1および仮説2に対して、4.3節の実験結果をもとにそれぞれの検証を行う。

4.4.1 仮説1の検証

仮説1では、質問文の再生成手法がTM 形成に資する情報収集および対話体験の質の向上に寄与するかどうかを検討した。

対話体験については、アンケート2の6つの設問すべてにおいて提案手法とベースラインの間に統計的に

有意な差は見られなかった（表3）。情報収集効果については、再生成プロセスを経た24件の回答の定性分析を実施した結果、有用情報を含む割合が50%（12/24件）となり、それ以外の56%（56/100件）と比較して改善は確認されなかった。

これらの結果から、仮説1は支持されなかった。

4.4.2 仮説2の検証

仮説2では、開発した情報共有システムの継続的利用によってTM形成が促進されるかを検証した。アンケート1の結果（表4）から、実験前、中、後のいずれの段階においてもTM形成に関するスコアは上昇しており、特に実験前後の比較においては統計的に有意な向上（設問1： $p = 0.0312$ ）が確認された。このことから、本システムが組織内のTM形成を段階的に促進する効果を持つことが実証され、仮説2は支持されたと結論付けられる。

4.5 考察

4.5.1 対話システムとしての効果

対話システムの質に関する仮説1については、統計的な有効性を示すことはできなかった。提案手法の有効性を統計的に示すことができなかった要因として、以下が考えられる：

1. 初期生成文の品質：最初に生成された発話が高品質であると判定された割合が57.8%と高く、再生成が不要なケースが多かった。
2. 認知的差異の閾値：初めからある程度整った質問文が生成されることで、改善が微細なニュアンスレベルにとどまり、被験者の認知的差異の閾値を超えなかった可能性がある。
3. 再生成プロセスで修正困難な根本的問題：対話ログ、フィードバックからシステムには以下のような根本的な問題が存在することが判明した。具体的には、時間認識の精度不足（先週の会議を今日の予定として言及）、情報の帰属混同（利用者自身の行動と他者の行動の区別不能）などが挙げられる。これらの問題を提案手法の評価・再生成プロセスでは解決できなかったため、アンケート2の設問1（自然性）、設問4（状況適合性）、設問5（回答しやすさ）、設問6（心理的負担）といった項目で統計的有意差に至らなかったと考えられる。

4.5.2 TM形成システムとしての効果

一方、TM形成に関しては本システムの有効性が示され、仮説2が実証された。これは、対話を通じて収集された情報とその共有により、メンバー間の相互理解が深まったことを示唆している。

ただし、システム設計上の課題として、「チャンネル＝プロジェクト」という前提に基づいた情報収集の限界が明らかになった。実際の組織運用では新規プロジェクト立ち上げ時に必ずしもSlackチャンネルが作成されるわけではなく、この仕様によりシステムが組織の動的な状況変化に十分追従できない場合があった。この問題を解決し、より包括的な組織情報の収集が可能になれば、レポートの精度向上やTM形成効果のさらなる促進が期待される。

現時点でも、システムから自動生成されるレポートに対する評価（設問5～7）は全て平均5.5であった。これは、参加者がレポートを自身および他者の活動理解に役立つものと感じていたことがうかがえる結果だと考えられる。この結果はユーザビリティだけでなく、実際の情報共有支援としてのシステムの価値を補足する結果であると考えられる。

4.5.3 質問応答プロセスの副次的効果

仮説1については支持されなかったものの、「提案手法を用いた期間の向上幅 \geq ベースライン期間の向上幅」という結果が見られ、傾向としては提案手法の有効性が示唆された。同一のレポートを全参加者が共有していたにも関わらずこの傾向が見られたことは、質問応答プロセス自体がTM認識能力の向上に寄与した可能性を示唆している。すなわち、質の高い質問に継続的に答える経験により、組織内の活動や関係性に対する注意の向け方や認知的枠組みが改善され、同じ情報からもより多くのTM関連要素を抽出できるようになったと考えられる。

5 結論

5.1 まとめ

本研究では、TM形成支援システムにおける質問文生成手法に着目し、TM形成、心理的・認知的負荷、自己効力感、文脈適合性の4つの観点から質問文を評価・再生成する手法を提案した。7名のベンチャー組織メンバーを対象とした実験により、以下の知見が得られた。

システム全体の効果：継続的な利用により組織内のTM形成が段階的に促進され、特にプロジェクト・分野の把握において統計的に有意な向上（ $p = 0.0312$ ）が

確認された。メンバー間の連携理解や自身の役割理解においても向上傾向が示された。

質問生成手法の効果：提案手法はシステムの自然性や親しみやすさにおいて数値的向上を示したが、統計的有意差は認められなかった。ただし、提案手法期間において TM 向上幅が大きくなる傾向が全設間で観察され、質問応答プロセスがユーザの認知能力向上に寄与した可能性が示唆された。

これらの結果から、本システムが組織内の知識共有と相互理解の促進に効果を持つことが実証された。特に、質問生成手法の改善は統計的には限定的であったものの、TM 形成という本来の目的においては一定の貢献を示すことが確認された。

5.2 今後の課題

本研究の成果を踏まえ、主に以下の課題が今後の研究において取り組むべき点として挙げられる：

対話制御と基盤技術の改善：本実験では対話が1回のやり取りで終了してしまう現象が発生し、対話制御に問題があることが明らかになった。加えて、時間認識の精度不足や情報の帰属混同などの根本的な問題が確認された。これらの問題に対して、プロンプト設計やシステム基盤の改善が必要である。

発話評価の精度向上：現在の提案手法における評価モジュールでは、組織と関係の薄い情報（例：個人の研究進捗など）を引き出す質問を適切に低評価できない場合があった。これは前述の情報不足とも関連しており、システムが組織の業務範囲や関心領域をより正確に把握することで改善される可能性がある。また、上記の問題に対しても適切に低評価を与えられるようなプロンプトの改善が必要であると考えられる。

組織情報の網羅的収集と実情把握：本研究では Slack データを主要な情報源としたが、これだけでは組織の現状を十分に網羅できていない。また、「チャンネル＝プロジェクト」という前提に基づいた情報収集は、新規プロジェクト立ち上げ時に必ずしも Slack チャンネルが作成されない実際の組織運用とシステムの認識の間に齟齬・乖離を生じさせた。この問題はシステムにそのような認識をさせたプロンプト設計に起因するものであり、組織の実態により即したプロンプトの改善が必要である。加えて、Slack 以外の多様な情報源を統合し、組織の変化により柔軟に対応できるスキームの構築が重要である。

また、上記の課題のほかに、より多くの被験者での実験、客観的評価指標の導入なども課題として挙げられる。これらに取り組むことで、TM 形成支援システムの実用性と効果をより高め、組織における知識共有と協働の質的向上に貢献できると考える。

謝辞

本研究の一部は、JST CREST (JPMJCR20D1)、篠原欣子記念財団、および科研費 (24K03052) の支援を受けた。

参考文献

- [1] Daniel M. Wegner. Transactive Memory: A Contemporary Analysis of the Group Mind, pages 185–208. Springer New York, New York, NY, 1987.
- [2] Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegrefe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, Yiming Yang, Shashank Gupta, Bodhisattwa Prasad Majumder, Katherine Hermann, Sean Welleck, Amir Yazdanbakhsh, and Peter Clark. Self-refine: Iterative refinement with self-feedback, 2023.
- [3] Saki Onuma and 沙樹 大沼. トランザクティブ・メモリー・システムと部門成果に関する実証研究. 商学研究科紀要, 82:1–20, 03 2016.
- [4] 克俊 宇佐美. 【教育学研究科高度教職実践専攻 学校課題解決研究報告書要旨】 tms の構築を通じた学校組織内の協働的關係 一学年を超えた教職員集団のつながりを深める研修の開発一. 教育デザイン研究, 16(3):63–66, 02 2025.
- [5] Kyle Lewis. Measuring transactive memory systems in the field: Scale development and validation. Journal of Applied Psychology, 88(4):587–604, 2003.
- [6] 角森 唯子 and 東中 竜一郎. パーソナライズ可能な雑談対話システム構築に向けた任意のユーザ情報を踏まえた発話コーパスの構築と分析. 自然言語処理, 31(4):1635–1664, 2024.
- [7] 東中 竜一郎, 高橋 哲朗, 堀内 颯太, 稲葉 通将, 佐藤 志貴, 船越 孝太郎, 小室 允人, 西川 寛之, 宇佐美 まゆみ, 港 隆史, 境 くりま, and 船山 智. 対話システムライブコンペティション 5. 人工知能学会研究会資料 言語・音声理解と対話処理研究会, 96:19, 2022.