

人格的記憶に基づく RAG の構築に向けた生成部の検討 Improving a Text Generation Method for Constructing Persona-based RAG

加藤 優一[‡] 市川 嘉裕[‡] 山口 智浩[‡]
Yuichi Kato Yoshihiro Ichikawa Tomohiro Yamaguchi

1. はじめに

近年、大規模言語モデル (LLM) を実用化する技術として検索拡張生成 (Retrieval Augmented Generation: RAG) [1] が利用されるようになってきた。RAG は、既存の LLM が持つ生成能力を活かしながら、特定のドメイン知識に即した応答の生成精度の向上を実現する。RAG において、LLM は外部知識から抽出した断片的な情報を一貫性ある文章へと成形する役割を担う。我々は、この外部知識を人格的記憶 (知識や経験を記述したデータ) に置き換えることで、特定の人格に即した応答生成が可能になるかという問いを立てた。併せて、特定の人格に即した応答生成を目的とするこのようなアプローチによる手法を人格的記憶に基づく RAG と呼称することとした。しかしながら、LLM による成形過程では LLM 内部に含まれる膨大な語群が参照されるため、人格的記憶に矛盾する応答や記憶に存在しない情報を用いた応答が生成される可能性があることが懸念される。これは通常の RAG ではより詳しい情報提供がなされるなどの点で有効に働くことにつながるが、特定の人格に即した応答としては不自然さに直結する問題となるため、これまでとは違った視点から詳しく調査する必要がある。

そこで本研究では、代表的なプロンプト設計手法の適用によって懸念される問題が抑止可能かどうかを実験的に検証することを目的とする。具体的には、(1)生成内容に制約条件を付加するプロンプト、(2)正例と負例を含めた Few-shot プロンプト、(3)適切な生成を期待する手順を指定した Chain-of-Thought プロンプトを設計し、ChatGPT に入力して生成された応答の質を比較する。本稿では、人格的記憶や検索部の厳密な設計はせず、生成部にのみ着目して調査を進める。

以上の実験を通じて、人格的記憶に基づく RAG の構築へ向けた問題点を整理し、より綿密な改善案の検討や人格的記憶に求められるデータ構造等の要件の明確化につながる知見を得ることを目指す。

2. 理論

2.1 検索拡張生成

事前学習済みの LLM を活用した検索拡張生成 (Retrieval Augmented Generation: RAG) は、生成 AI に外部情報源を組み合わせることで、外部知識に基づいた信頼性の高い応答を生成する手法である[2]。モデル自体に膨大な知識を内包させる必要がないため、効率的なモデル設計が可能となる。特定のドメインに特化した知識ベースを活用することで、多様なタスクに対応できるスケーラビリティも備えている。

一方で、RAG にはいくつかの課題も存在する[2]。出力の品質が検索の品質に大きく依存する点がその一つである。

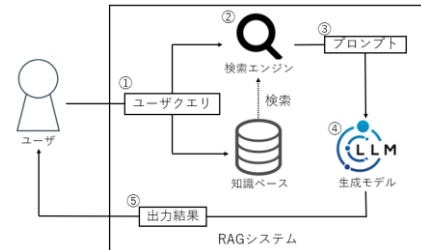


図 1 RAG の概略図

外部検索結果が不正確である場合や、関連性が低い場合には、生成内容にも悪影響を及ぼす。また、検索と生成を組み合わせる仕組み上、計算負荷が増加し、処理コストが高くなる傾向がある。さらに、検索結果とモデルの生成内容が矛盾することで一貫性の問題が生じる可能性や、外部情報源の可用性や信頼性に依存する点も課題として挙げられる。

RAG は図 1 に示すように構成される。

- ① ユーザクエリの生成：
ユーザからの質問を受け付け、次の検索プロセスに適した形式のユーザクエリを作成する。
- ② 関連知識の検索：
検索エンジンを用いて、ユーザクエリに関連する外部知識を知識ベースから検索する。
- ③ プロンプトの作成：
検索結果として得られた外部知識を基に、ユーザの質問と関連知識を組み合わせた形で LLM への入力となるプロンプトを作成する。
- ④ モデルによる生成：
作成したプロンプトを LLM に入力し、回答を生成する。
- ⑤ 回答の抽出と出力：
生成された結果の中から、ユーザの質問に対する回答部分を抜き出し、最終的な出力としてユーザへ提示する。

2.2 プロンプト設計手法

ここでは、LLM の入力として利用される代表的なプロンプト設計手法について述べる。

2.2.1 深津式プロンプト

深津式プロンプトとは、note 株式会社の CXO である深津貴氏が考案した、生成 AI を効果的に活用するためのプロンプト設計手法である[3]。この手法は、ChatGPT をはじめとする生成 AI への指示をより明確かつ具体的にすることで、ユーザの期待に沿った出力を得ることを目的としている。深津式プロンプトは、学術的な裏付けが明示されているわけではないものの、実務において品質向上に寄与すると評価されている[4]。その特徴として、単に指示を与えるだけでなく、AI との対話プロセス全体を効率的に設計

[‡] 奈良工業高等専門学校 National Institute of Technology (KOSEN), Nara College

することで、出力結果の精度や有用性を高める点が挙げられる。本手法は、生成 AI の性能を最大限に引き出し、利用者が意図する情報や成果物を得るための体系的なアプローチとして注目されている。

深津式プロンプトは以下の特徴を持つ。

1. 役割を明確にする
2. 入力から出力を明確にする
3. 何を出力するかを明確にする
4. マークアップ言語を用いて本文以外を明確にする
5. 命令を箇条書きで明快にする
6. 条件を指定して出力結果をコントロールする

2.2.2 Few-shot プロンプティング

Few-shot プロンプティングは、少数の例を用いて大規模言語モデルに新たなタスクを学習させる手法である。この手法は、特にデータ収集が困難なタスクやリソース制約のある状況で有用である。Brown らの研究[5]では、GPT-3 を用いた Few-shot プロンプティングの有効性が示されている。具体的には、数個の例を入力プロンプトとして与えるだけで、モデルが多様なタスクを高い精度で実行できることを実証した。

さらに、日本国内においても Few-shot プロンプティングを活用した研究が活発に進められている。例えば、ChatGPT を用いて特定の人物の人格を再現する試みが報告されている[6]。この研究では、わずか 10 個の発言例を用いることで、追加の訓練を行うことなく人格の再現性を向上させる手法が提案されている。このような事例は、Few-shot プロンプティングが幅広い応用可能性を持つことを示唆している。

2.2.3 Chain-of-Thought プロンプティング

Wei らは Chain-of-Thought プロンプティングを提案し、複雑な推論タスクにおけるモデル性能の大幅な向上を報告している[7]。Chain-of-Thought プロンプティングとは、解答に至るプロセスを連鎖的に展開する形式でのプロンプト設計をすることであり、特に論理的な推論を要するタスクにおいて有効であることが確認されている。

3. 実験

3.1 概要

本実験では、RAG におけるデータの検索プロセスは対象とせず、質問に関連するとして抽出された人格的記憶のセットを用いて LLM に回答を生成させる部分に着目して検証を行う。具体的には、一つの会話シチュエーションを設定し、特別な工夫をしないプロンプトを用いて出力を行うことで、RAG において起こる問題を明確にする。次に、代表的なプロンプト設計手法を参考に設計した三つのプロンプトを用いて出力を行い、出力内容を比較する。これにより、各手法が当該問題に与える効果を検証する。

3.2 設定

3.2.1 会話シチュエーション

本実験で使用する会話シチュエーションについて述べる。具体的には、RAG に入力する質問内容と RAG が参照する人格的記憶の内容の二点を示しその意図を付記する。

RAG に入力する質問は「今度の瀬戸内旅行どこに行きたい?」とした。質問の選定にあたり定めた要件は次の三点通りである。

- ・ 個人の知識や記憶によって回答が変わるもの
 - ・ 主観的な意見を問うもの
 - ・ 現実の事象と深く関連しているものであること
- 一、二点目は、人格的記憶に基づく RAG が特定の人格に即した応答を生成できるかどうかを確認する点で重要である。三点目は、人格的記憶にない LLM 内部の情報と結びつく可能性が高く、問題が生じやすいシチュエーションとなることを想定して定めたものである。

RAG が参照する人格的記憶は、実験のために簡素化したものを扱うこととした。具体的には、RAG に入力された上記の質問に対して、人格的記憶の中から関連する情報が抜き出されることを想定し、関連する情報を箇条書きでリスト化したものとする。本実験では次のように設定した。

- ・ 趣味はサイクリングである
- ・ 海沿いの風景を楽しむのが好き
- ・ 旅先では地元の自然や景観を重視する
- ・ 西日本に住んでいる
- ・ 小学生の頃、香川県でうどん作り体験をした
- ・ 出張で岡山県を訪れたが、観光はしていない
- ・ 飛行機を使った長距離旅行をよくする
- ・ 山登りが好き
- ・ 九州地方に親戚がいる
- ・ バスツアーを利用することが多い
- ・ 旅行は好き

本稿では RAG の検索プロセスや情報源の形式について定めていないという前提で、質問の文中にある単語に強く関係する情報と弱く関係する情報が多様に抽出されたことを想定している。注意として、瀬戸内周辺の詳しい地名や名物についての知識を持たないことがあげられる。

以上の設定より、筆者が期待する回答は「瀬戸内海沿いをサイクリングしたい」「岡山観光リベンジしたい」「またうどん作りをしたい」など、記憶を基にした抽象的に留めておくものである。反対に「しまなみ海道でサイクリングしたい」「倉敷の街を練り歩きたい」「小豆島が人気だよね」など、記憶にない詳しい地名や名物についての知識を利用する回答は期待しておらず、当初の問題が生じているケースとして評価する。

3.3 プロンプト

実験で扱うプロンプトについて述べる。本稿では特別な工夫をしないプロンプトをデフォルトプロンプト、代表的なプロンプト設計手法を参考に設計した三つのプロンプトをそれぞれ制約付加プロンプト、Few-shot プロンプト、CoT プロンプトと称する。

表 1 にデフォルトプロンプトを示す。これは RAG システムを用いて会話応答生成を行う上での最低限の機能のみ備えたプロンプトである。実験の第一段階では、これをベースに起きうる問題について考察する。

表 2 に制約付加プロンプトの一部を示す。LLM への入力はデフォルトプロンプトの末尾に表 2 の文を追加したものを利用する。このプロンプトは回答で想定される問題を直接禁止するものであり、特に人格的記憶との矛盾と人格的記憶に無い情報の使用を禁じることを明記している。

表 3 に Few-shot プロンプトの一部を示す。ここでは、デフォルトプロンプトの前に、類似タスクの指示とその回答として良い例と悪い例、悪い例である理由を示している。

表 1 デフォルトプロンプト

<p>#指示 あなたは会話をしています 以下の人格的記憶を持った人物になりきって相手の発言に返答してください</p> <p>#人格的記憶 (人格的記憶の箇条書きリスト)</p> <p>#相手の発言 「今度の瀬戸内旅行どこに行きたい？」</p> <p>#出力形式 返答内容のみを出力</p>
--

表 2 制約付加プロンプト（一部抜粋）

<p>#制約条件</p> <ul style="list-style-type: none"> ・人格的記憶と矛盾する内容を含めないこと ・人格的記憶に無い情報は応答内容に含めないこと
--

その後、デフォルトプロンプトによる指示により、元タスクに対する回答を求める形式になっている。類似タスクとして北海道旅行を相談するシチュエーションと、それに關する人格的情報を設定した。良い回答の例は同じプロンプトを用いて実際に LLM によって出力された回答を実験者が推敲したものである。なお、類似タスクの人格的記憶と元タスクの人格的記憶が同一人物のものであるかの指定は明記していない。実験者の意図はタスクが異なるため別々の人物であることを想定しているが、一つのプロンプトに記載しているため LLM が混同することは想定される。

表 4 に CoT プロンプトの一部を示す。LLM の入力にはデフォルトプロンプトの「#出力形式」以降を表 4 の文に置き換えたものを利用する。このプロンプトでは、応答生成の手順を三つのステップに細分化し、さらにステップごとに出力させる。多数の人格的記憶から一つに絞ってフォーカスさせることによって、漠然と人格的記憶に矛盾する応答や記憶に存在しない情報を用いた応答を生成することが抑制されることを期待する。

3.4 評価基準

本実験では、付加した人格的記憶と矛盾する内容をエラーとして扱う。それらの発生頻度をもとに回答の質を評価する。問題点として扱う条件は、人格的記憶に存在しない詳細な情報（知らないはずの物事）について話してしまうことである。

3.5 実験手順

本実験は、以下の手順に従って進める。

- 3.2 で示した会話シチュエーションと人格的記憶を、3.3 で示した各種プロンプトに適用し、ChatGPT4-o (Free プラン, 2025 年 6 月実行) に入力して回答を出力させる。これをプロンプトごとに 20 回ずつ行う。
- 出力結果に対し、3.4 で設定した条件に当てはまる内容があればエラーとし、その発生回数を記録する。

表 3 Few-shot プロンプト（一部抜粋）

<p>#例</p> <p>##指示 あなたは会話をしています 以下の人格的記憶を持った人物になりきって相手の発言に返答してください</p> <p>##相手の発言 「今度の北海道旅行どこ回ろうか？」</p> <p>##人格的記憶</p> <ul style="list-style-type: none"> ・趣味はドライブである ・海鮮料理が好き ・旅先では地元の名物料理を食べることが楽しみ ・西日本に住んでいる ・近頃、静岡県で浜松餃子を食べに行った ・インターンで東京を訪れた ・車を使った長距離旅行をよくする ・友達との会話が好き ・東北地方に親戚がいる ・旅行が好き <p>##良い回答 北海道って言ったら、やっぱり海鮮でしょ！どのエリア回るかにもよるけど、海鮮目当てでルート組むのも全然アリやと思う！移動手段どうする？レンタカーならドライブも楽しめるよ！</p> <p>##悪い回答 いいね北海道！広いからどこ行か悩むよね？。札幌は外せないとして、函館で朝市の海鮮丼とか最高だよ。あとは美瑛・富良野の景色も癒されるし、旭川でラーメン食べるのもアリ。時間あれば道東の知床とかも行ってみたいなー。移動手段どうする？レンタカーならドライブも楽しめるよ！</p> <p>##悪い理由 函館の朝市、美瑛・富良野の景色、旭川でラーメンなど、人格的記憶に無い知識が含まれている。 「なりきって答える」という指示に対して、人格的記憶に無い知識を用いるのは適さない。</p> <p>#指示 (中略)</p> <p>#良い回答</p>

表 4 CoT プロンプト（一部抜粋）

<p>#出力形式 下の手順に沿って出力</p> <p>#手順</p> <ol style="list-style-type: none"> 1. 相手の発言と最も関係のある人格的記憶を一つだけ抜き出す 2. 抜き出した根拠を考える 3. 抜き出した人格的記憶と根拠に基づいて返答を出力する

4. 実験結果

表 5 にエラーの発生回数を示す。

プロンプト	発生回数 (20 回中)
デフォルトプロンプト	20
制約付加プロンプト	8
Few-shot プロンプト	5
CoT プロンプト	10

全体を通して、人格的記憶にない「しまなみ街道」「小豆島」など具体的な地名に関する情報が多々見られた。

デフォルトプロンプトでの出力例として、『いいね、瀬戸内旅行！せっかくだから海沿いを自転車で走れる場所がいいな。しまなみ海道とかどう？景色がほんとに綺麗だし、橋を渡るたびに島ごとの雰囲気も楽しめるよ。それか、昔うどん作り体験した香川でまた本場の讃岐うどん巡りってのもアリかも。』といったものがあり、「自転車」や「うどん作り体験」など、趣味や経験に基づいた自然さを感じる応答である一方で、「しまなみ海道」を未知の情報を含むというエラーが発生したとしてカウントしている。デフォルトプロンプトによる出力の 20 回すべてで「しまなみ海道」が登場する結果となった。

制約付加プロンプトでは、出力例として、『うーん、せっかくなら海沿いの景色が楽しめるところがいいな。前に香川でうどん作り体験したのが楽しかったから、また香川に行ったら本場のうどんを味わいつつ、瀬戸内海沿いをサイクリングするのもいいかも。天気良ければ、島に渡って景観を楽しむのもいいね。』のように、エラーとカウントしない例が増え四つの条件の中では二番目に良い結果となった。しかしながら、デフォルトプロンプトと同様のエラーが発生することも見られた。

Few-shot プロンプトの出力例としては『瀬戸内って聞くと、やっぱり海沿いの道を自転車で走ってみたいかな？。潮風感じながらのサイクリングとか最高やろ！昔、香川でうどん作りしたの思い出すし、また行ってみるのもアリかも。今回は自然とか景色重視でルート考えてみる？』のようなものがあり、文章量は変わらないものの、人格的記憶にない知識を避ける例が大幅に増えた。

CoT プロンプトでは、人格的記憶にない情報の使用を明示的に禁止していないにもかかわらず、デフォルトプロンプトに比べてエラーの発生頻度が大きく抑えられた。出力例をみると、『うーん、香川にまた行ってみたいなあ。小学生の頃にうどん作り体験をしたことがあって、あのときの楽しさが今でも印象に残ってるんだよね。今度は大人の目線で、もう少しゆっくりと街を巡ったり、海沿いの風景も楽しみたいな。』のように、手順通りに一つの人格的情報を軸として、それを膨らませて文章を生成した内容の出力がなされていることも確認できた。

5. 考察

5.1 プロンプト間の比較

人格的記憶に無い情報が返答に含まれる問題に対し最も効果的だったのは Few-shot プロンプトである。しかし、Few-shot として例示した会話シチュエーションと人格的記

憶が、本タスクと非常に類似していることが要因であることが考えられる。その裏付けとして、生成された応答を確認すると明らかに正例に似た文章になっている。例えば、良い例の最後に『移動手段どうする？レンタカーならドライブも楽しめるよ！』という移動手段についての言及があるが、Few-shot の結果でも『バスツアーとかも組み合わせたら、移動も楽で景色も満喫できそうやし！』など、多くの事例で移動手段の言及がある。このことから、シチュエーションや人格的記憶の違う場面に対して効力を発揮できるかという汎用性に疑問が残る。

人格的記憶に無い情報の使用の抑制について、制約付加プロンプトは Few-shot プロンプトより効果が低かった。考えられる原因として、使用して良い情報かどうかの線引きが曖昧な点が挙げられる。断片的な情報のまとめ上げをするには、行間を補う必要があり、それをしない場合、たどたどしく不自然な文章になってしまう。このことから、一般性の高い知識や現在の話題と関連性の高い知識が意図せず表出してしまったのではないかと考えられる。これらの使用しても良い情報の優先度付けが人間（今回は実験者の主観）と異なり、問題として判定されたといえる。ただし、Few-shot プロンプトと比較して詳しい状況を指定することなくある程度の効果が認められたことから、汎用性の高い手法であることが評価できる。

CoT プロンプトは他のプロンプトよりもエラーを防止する効果は低かったが、出力過程において大きな知見が得られた。具体的には、20 回の生成の中で、CoT の手順 1 で選ばれた人格的記憶は合計で三種類であった。そこで、「旅先では地元の自然や景観を重視する」と「海沿いの風景を楽しむのが好き」が選ばれた際は必ずエラーが起きていた。一方で、「小学生の時、香川県でうどん作り体験をした」が選ばれた際はエラーが起きなかった。この原因を詳細に分析することで、エラーの起こりにくい応答生成につながる有効な知見が得られると考える。

5.2 返答の文章量について

各プロンプトが生成した応答の文字数を調べた結果を表 6 に示す。

プロンプト	平均	標準偏差
デフォルトプロンプト	110.80	12.10
制約付加プロンプト	104.95	16.40
Few-shot プロンプト	142.55	17.75
CoT プロンプト	103.00	13.70

デフォルトプロンプト以外のプロンプトはデフォルトプロンプトに比べてなんらかの制約を加えられていると捉えることができる。制約付加プロンプトや CoT プロンプトの応答文の文字数が減少したのは、表現が制限されることによると考えられる。一方で、Few-shot プロンプトは文字数が増加する結果となった。5.1 で述べたように、Few-shot プロンプトは良い回答の例をなぞった生成をしようとするが故に、デフォルトプロンプトよりも文章が長くなったと考えられる。

5.3 文脈の不整合について

本実験では、「行きたい場所」に関する質問を提示し、LLM の応答傾向を観察した。その結果、『瀬戸内って言ったら、やっぱり海沿いの風景が気持ちよさそうやね！個人的にはサイクリングで海を眺めながら回るの、憧れるわ〜。』など、質問に対して直接的な地名や具体的な場所を挙げるのではなく、情緒的あるいは主観的な印象や願望を交えた応答が多数確認された。

一見すると、これらの応答は質問への明示的な回答を回避しているように見えるが、会話応答の観点からは自然かつ妥当な返答であり、会話の流暢さや多様性を高める効果が認められる。このような傾向は、会話応答システムにおいて単に情報を返すのではなく、対話文脈に即した応答生成がなされていることを示唆している。すなわち、本システムはユーザの問いに対して一義的な解を返すのみならず、会話の流れや情緒を考慮した応答を生成しており、対話エージェントとしての自然性を備えていることが確認された。

5.4 主観的表現の扱いについて

本実験においては、『(前略)あとは、香川にも寄って、久しぶりにうどん打ち体験してみたいな。あの頃と比べて、今なら上手にできるかもって思っ。』といった、人格的記憶に含まれない主観的表現が一部の返答に見られた。これらの要素は人間同士の自然な会話において不可欠なものであり、本研究ではこれをエラーとは見なさず、分析対象に含める判断を行った。しかしながら、こうした主観的応答が、人格的記憶によるものではなく、大規模言語モデル(LLM)の学習データに由来して生成されている場合、応答の一貫性や人格の再現性を損なう可能性がある点は看過できない。ゆえに、将来的には、主観や感情といった内面的情報もプロンプト内に明示的に含める必要があると考えられる。

ただし、こうした情報は、現在の知識や経験といった記憶要素とは性質や表現の方向性が大きく異なるため、単純に既存のプロンプト形式に統合することは困難である。したがって、主観的要素を適切に記述・制御可能な新たなプロンプトフォーマットの検討が求められる。

5.5 実験の改善点

本実験では、RAG を用いた会話応答生成における課題の明確化および各種プロンプト手法の特性を検証することができた。しかしながら、いくつかの点で実験設計に改善の余地が認められる。

第一に、本実験で設定したシチュエーションは研究者が人為的に設計したものであり、実際の利用環境下における自然な対話状況とは乖離している可能性がある。そのため、今後は実際のユーザとのインタラクションに基づく評価や、現実的な利用ケースを想定した対話ログの活用が求められる。

第二に、本実験は単一のシチュエーションについてのケーススタディに留まっており、各プロンプト手法の特性を包括的に検証したとは言い難い。特に、特定の状況下においてのみ発現するようなプロンプト特有の挙動については、本実験では十分に明らかにされていない。したがって、今後は多様なシナリオに基づいた検証を通じて、より一般化可能な評価を行う必要がある。

第三に、Few-shot プロンプトにおいて提示した会話例が、実験に用いたシチュエーションと過度に類似していたことは、生成された応答に対するバイアスを引き起こす要因となり得る。このような事前知識の過学習を避けるため、より多様かつ一般化された会話例を用いる設計が必要である。

第四に、CoT プロンプトに関しては、本実験では単一の推論手順に基づくプロンプトを適用したが、他にも複数のバリエーションが考えられる。したがって、今後は異なる推論戦略を比較検討し、プロンプト構造の最適化を図るべきである。

最後に、本研究では ChatGPT-4o (Free プラン) を用いて実験を実施したが、同モデルは質問応答タスクに最適化されたファインチューニングがなされており、この特性が会話応答の結果に影響を及ぼした可能性がある。そのため、その他異なる特性を有する言語モデルを用いた比較検証を行い、人格再現に適したモデル選定を行うことが求められる。

6. おわりに

6.1 まとめ

本研究では、「特定の人格に沿った会話応答生成」を目的とし、人格的記憶をプロンプトに付加して会話応答を生成する RAG システムを仮定した。その上で、人格的記憶と矛盾する応答、特に記憶に存在しない詳細な情報を含む応答をエラーと定義し、各種プロンプト設計手法が当該エラーの抑制に与える影響について検証を行った。

実験の結果、Few-shot プロンプトが最も高いエラー抑制効果を示した。しかしながら、これはプロンプト内で例示された会話シチュエーションと人格的記憶がタスク内容と極めて類似していたためと推察され、得られた効果の汎用性には課題が残ると考えられる。

制約付加プロンプトにおいては、応答に人格的記憶にない情報を含めないよう明示的に制約を設けたにもかかわらず、Few-shot プロンプトに比してエラー抑制効果は低かった。この結果については、使用してよい情報の範囲が曖昧であり、モデル側が判断に迷った可能性が要因として考えられる。

一方、CoT プロンプトでは、エラー抑制効果は他の手法よりも低かったが、応答生成の過程において選択された人格的記憶が出力に及ぼす影響を段階的に分析できた点に意義がある。本手法は応答生成の透明性と分析可能性を高める観点から、今後の発展が期待される。

これらの結果を踏まえ、プロンプト設計と人格的記憶の付加方法を適切に調整することで、人格情報に基づく応答生成の精度向上が期待できる。

6.2 今後の課題

本研究では、RAG システムにおける出力生成部に焦点を当て、人格的記憶の有無が生成出力に与える影響を検証した。しかし、実際の対話システムにおいては、ユーザからの入力文に対して関連する記憶を検索・抽出する検索部の性能も重要である。したがって、今後は本研究で得られた知見をもとに、検索部における記憶検索手法の精度およびその出力生成への影響についても検証を行う必要がある。

また、本研究では「特定の人格に沿った会話応答生成」を目的とし、主に経験や知識といった明示的な人格的記憶

の反映に着目した。しかし、人格を構成する要素はこれに限られず、感情傾向や価値観、言語表現の癖といった非明示的・暗黙的な情報も人格再現において重要な役割を果たす。今後は、これらの要素を生成出力に反映可能とするプロンプト設計や制御手法の開発、およびその評価指標の確立を通じて、より精緻な人格再現の実現を目指す必要がある。

参考文献

- [1] Patrick Lewis, Ethan Perez, Aleksandra Piktus, “Retrieval-augmented generation for knowledge-intensive nlp tasks”, arXiv (2020).
- [2] Yunfan Gao, Yun Xiong, Xinyu Gao, “Retrieval-augmented generation for large language models: A survey”, arXiv (2023).
- [3] MiraLab AI, “Chatgpt の深津式プロンプトとは？コピペで使える例文やテンプレート”, https://miralab.co.jp/media/fukatsu_prompt/ (閲覧日 2024-1-21) (2024).
- [4] Sondas Mahmoud Bsharat, Aidar Myrzakhan, “Principled instructions are all you need for questioning llama-1/2, gpt 3.5/4”, arXiv (2023).
- [5] Tom B. Brown, Benjamin Mann, “Language models are few-shot learners”, arXiv (2020).
- [6] 坪田 侑華, 狩野 芳伸, “10 の発言例で構成する few-shot prompting により人格を再現した llm 文生成”, 人工知能学会全国大会論文集, 38, (2024).
- [7] Jason Wei, Xuezhi Wang, “Chain-of-thought prompting elicits reasoning in large language models”, arXiv (2022).