

個人の経歴情報に基づいた質問生成機能の検討

Study on Question Generation Function Based on Personal Background Information

伊藤 奈桜[†] 間瀬 正啓 田中 毅 梁 宇シン[‡]
Nao Ito Masayoshi Mase Takeshi Tanaka Yuxin Liang

1. はじめに

「面談」は現代社会において広く使用されるコミュニケーション手段である。インタビューやカウンセリングなど多様な場面で実施されるものであるが、特に人事採用における面談は、企業が優秀な人材を獲得するために応募者の経歴や能力を測る場として非常に重要である。近年はオンラインによる採用面談を実施する企業も多く、人材発掘の効率化に関して需要が高まっている。

このような対話を通じて情報を引き出す場面において、大規模言語モデル（Large Language Model, LLM）の利用が期待されている。インタビューやカウンセリングにおいて聞き手となる立場には、特定の方針に従って相手の発言を促すために適切な振る舞いが求められる。近年は LLM においてプロンプトエンジニアリングを活用し、特定の目的に沿って発言を引き出す対話システムの構築が提案されているが、主なアプローチは対話中の文脈に基づき発言内容を生成するものである。しかし、情報を引き出すことが目的の対話においては、前後の文脈だけでなく、発言者の背景情報や特性などに基づき会話が行われることもある。LLM を活用して、対象者の経歴や発言内容に基づき適切な発言や質問文を生成する具体的手法については十分に検討されていない。

本研究では、LLM を用いて、対象者の背景情報に基づき深堀質問を自動生成する手法を提案し、その有効性を検証することを目的とする。特に、対話場面として人事面接を想定したシナリオを対象とし、生成された質問の妥当性を評価する。

2. 関連研究

対話における質問生成に関する先行研究として、土屋ら[1]はカウンセリングで行われる動機づけ面接において、クライアント自身の変化に向けた発言を促進させる面接会話エージェントの構築方法を提案した。この先行研究では、対話の中でクライアント自身が問題の全体像を把握し解決のための糸口を見つけるために、LLM の一つである対話型 AI の ChatGPT[2]を用い、プロンプトエンジニアリングによってクライアントの発言内容を要約しさらなる発言を促進する方法が提案されている。これはスタンスとしては対話で得られた情報の要約を提示しクライアントの発言を待つという受動的なものであり、引き出したい観点を的確に聞き出すための質問生成方法については論じられていない。また、面談において相手を掘り下げる質問を動的に生成する先行研究として、ニューラルネットワークを活用したモデル[3]が提案されている。学習データが十分にあれば

各種の対話場面において効果的な掘り下げ質問を生成可能と考えられるが、一方で学習データへの依存性があり、対話の場面や目的に応じた調整が必要であると考えられる。さらに、就職面接において面接官が効きたい観点と、これまでの応募者の発言内容からさらに内容を掘り下げる対話システム[4]も提案されている。これはベースとする質問とそれに対応するチェック項目を予め設定し、質問に対する回答が充足しているか否かを判定し、その結果に応じて掘り下げ質問を生成するものである。これは対話の目的が特定の情報収集である場合に非常に高い有効性を示すと考えられるが、掘り下げ方を逐次設定する必要があるため、採用方針が異なる企業間ではそれぞれ調整が必要と考えられる。本研究では、先行研究[4]と同様に聞きたい事項が決まっているものの、どのように掘り下げするかについては自動かつ動的に制御する手段として、対象者の背景情報を活用する方法について検討する。

3. 提案手法

本研究では、対象者の背景情報として経歴を活用し、深く対話を進めるための「深堀質問」を自動生成する手法を設計した。本研究では人事採用面接の場面において人間と AI が 1 対 1 で対話することを想定する。以降、対話に参加する人間を応募者、AI を面談官 AI と表現する。

3.1 使用データ

表 1 に示す通り、本手法が用いるデータは応募者の履歴書と既存の質問文の 2 種類である。図 1 に応募者の履歴書の例を示す。面談前に応募者が自身の経歴を記載した履歴書を提出することを想定している。また、既存の質問文は例えば「過去にリーダーシップを発揮した経験はありますか？」等、一般的な人事採用で用いられる質問文のリストである。各企業が採用方針に則した質問リストを保持していることを想定している。

表 1 深堀質問生成システムの入力データ

データ	説明
応募者の履歴書	応募者の経歴データ。過去に行った業務内容が記載されている。
既存の質問リスト	面接官が所持している質問文のリスト。人事面接でよく用いられる一般的な質問内容。

[†] 日立製作所 研究開発グループ[‡] Research & Development Group, Hitachi Ltd.

期間	所属	業務の名称	業務の担当内容・実績	開発環境や使用スキル	役割/規模	重要な経験やアピールポイント
2023年4月～現在	(株)日立製作所	AI対話サービスの開発	自動対話システムの研究	Python, OpenAI	サブリーダー/5名	特許1件を出願済み
...

図1 応募者の履歴書の例

3.2 深堀質問生成アルゴリズムの構築

本研究では、大規模言語モデル（LLM）の ChatGPT を用いて質問生成システムの検討を行った。図2に提案手法におけるデータ処理の流れを示す。提案手法は大きく2段階の処理から構成される。

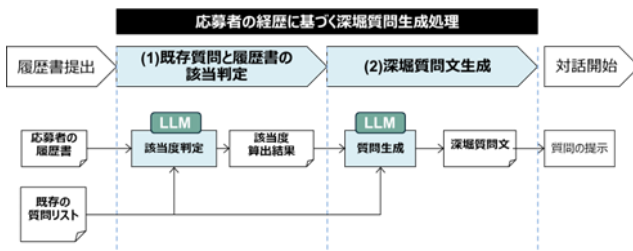


図2 提案手法による質問生成のフロー

3.2.1 既存質問文と履歴書の該当判定

1段階目の処理では ChatGPT を用いて応募者の履歴所と既存の質問文を照合し、面接官が問うことになる応募者の経験やスキルを応募者が持ち合わせているか事前に判定した。具体的には質問文に対して回答となりうる情報が履歴書内に存在するかを示すため該当度という指標を作成した。該当度の判定基準は表2のように策定した。プロンプト内で該当度について説明し、応募者の履歴所と既存の質問文の該当度を ChatGPT に判定させた。さらに、プロンプト内で、該当度2または3と回答した場合はその根拠となるテキスト情報を履歴書から抽出し、「経歴証拠」として提示するよう指示した。

表2 質問文と履歴書の該当度

該当度	質問文と履歴書の関連性
1	既存の質問文に関連する情報が履歴書に無い
2	既存の質問文に関連する情報が履歴書にあるが、確実に経験しているとは言えない
3	既存の質問文に関連する情報が履歴書にあり、実際に経験している可能性が高い

3.2.2 深堀質問文の生成

2段階目の処理として、ChatGPTにより既存の質問文を履歴書の情報に基づいて具体化した。プロンプト内で、該当度が2と算出された場合、すなわち既存の質問文に対応する経験またはスキルが応募者にある可能性が高いと判断された場合、1段階目の処理で抽出された「経歴証拠」を

引用しながら既存の質問を問うようなアレンジを行うよう指示した。さらに、応募者からより明確な発言を得られるよう、採用面談の場面においてしばしば活用される STAR の観点をを用いて質問文を具体化した。STAR は面接官が具体的かつ効果的な質問を投げかけるためのノウハウであり、以下の4つの観点に基づく質問方法である。表3に4つの観点とその説明を示す。プロンプト内で、STAR 質問の4つの観点について説明し、アレンジされた質問文の後に1つ観点を選んで追加文を生成するよう指示した。

表3 質問生成に用いた STAR の観点

観点	説明
Situation	状況を問うための質問。例:「どのような環境でしたか」
Task	課題を問うための質問。例:「何が大変でしたか」
Action	行動を問うための質問。例:「どのように行動しましたか」
Result	結果を問う質問のための。例:「その結果どうなりましたか」

3.3 評価設計

人事面接を想定したシナリオを評価対象とし、提案手法によって生成された深堀質問文の妥当性を評価した。被検者は1年以上の社会人経験をもつ8名の会社員を対象とした。被験者は自身の履歴書を事前に提出し、Webブラウザ上でテキストベースの模擬面談に参加した。模擬面談では事前に設定された一般的な質問文と、提案手法によって生成された個人の経歴情報に基づく深堀質問が面接官 AI から提示され、被験者は各質問に対してチャット形式で回答を入力し、その後3ターンの対話を面接官 AI と行った。模擬面談後、被験者は質問の妥当性に関するアンケートに回答した。①適切性、②個人への寄り添い、③難易度という3つの観点に関して質問を提示し、被検者は1から5のいずれかの値で回答した。各観点に関し提示した質問文を表4に示す。適切性、個人への寄り添いについては回答値が高いほど、難易度については回答値が低いほど良好な回答となるように回答値を設定した。提案手法により生成された深堀質問と、事前に設定された一般的な質問文について、上記3点の評価結果を比較した。

表 4 アンケートの評価観点

観点	質問と回答形式
適切性	Q.実際の面談を想定した際、質問文ほどの程度適切だと感じましたか？(1:悪い 2:やや悪い 3:普通 4:やや良い 5:良い)
個人への寄り添い	Q.この質問はあなた個人の経歴に寄り添った質問でしたか？(1:全く思わない 2:思わない 3:わからない 4:そう思う 5:強くそう思う)
難易度	Q.この質問の難易度はどのくらいでしたか？(1:非常に簡単 2:簡単 3:普通 4:難しい 5:非常に難しい)

4. 結果

一般的な質問文と提案手法の深堀質問文それぞれに対する①適切性、②個人への寄り添い、③難易度それぞれの評価結果を図 3,4,5 に示す。図中に記された各スコアは 8 名の被検者の回答の平均値を示す。提案手法で生成された深堀質問文は、あらかじめ設定された一般的な質問文と比較して、適切性・寄り添いの度合いが高く、また回答の難易度が低いことが示された。

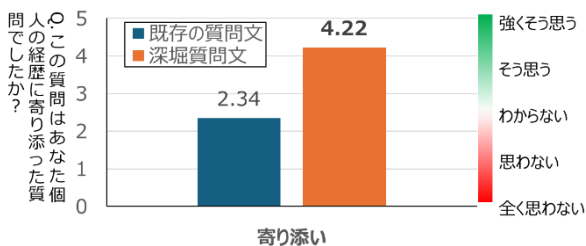


図 3 質問文に対する適切性の評価結果

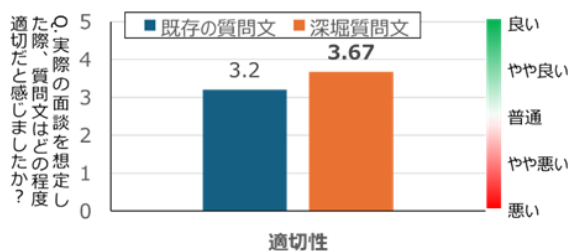


図 4 質問文に対する個人への寄り添いの評価結果

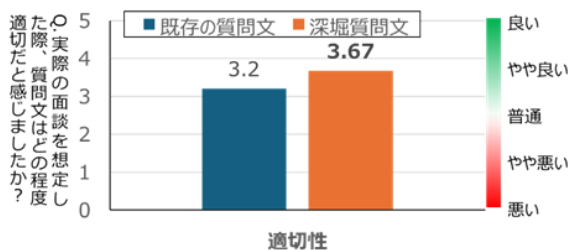


図 5 質問文に対する難易度の評価結果

5. 考察

5.1 提案手法の有効性と課題

本研究の結果は、LLM を用いることで、個人の経歴や特性といった背景情報に基づいた深堀質問を生成できる可能性を示唆するものである。各指標の改善幅を比べると、提案手法により適切性は 0.47、寄り添いは 1.88、難易度は 0.54 改善したことがわかる。寄り添いに関する回答値が最も大きく改善した理由については本手法が個人の経歴に言及する形で質問文を具体化したことがダイレクトに影響したためと考えられる。また、適切性に関して全体的には改善が見られたものの、一部の被検者からは「履歴書から引用された箇所が不自然に感じた」という感想を頂いた。例えば「自身の問題点を指摘された際、どのように対応しますか？」という旨の既存質問に対し、経歴情報として、ある被験者の履歴書からは「要件定義、システム基本設計」という実際の作業内容に関する情報が抽出された。このときの該当度は 2 であった。その結果、後段で生成されたのは、「{特定のサービス}の開発に関するご経験があるようですが、要件定義に関わる業務の経験を活かして、自身の問題点や悪いところに関する指摘を受け入れる方法や自己成長に活かす方法を教えてください。」という深堀質問であった。このように質問と抽出する経歴情報の関連性が低い場合、後段で生成される質問文も限定的で不自然な文となり、かえって被検者が回答に困ることが確認された。このような課題への対策案として、入出力データの細分化が挙げられる。提案手法の第 1 段階で行った該当度の算出と経歴証拠の引用に関して、判定ロジックの高精度化が今後の課題として考えられる。

5.2 個人情報の取扱いと質問生成に関する倫理的課題

本研究で提案した深堀質問生成手法においては、応募者の履歴書や経歴情報といった個人情報を LLM に入力する処理が含まれる。このような情報の取り扱いに際しては、データのプライバシー保護と倫理的配慮が重要である。具体的には、履歴書情報の利用に関して応募者の事前同意を得る仕組みが必要である。さらに、LLM を利用する際に、入力データが外部サーバに送信される場合のセキュリティ対策も必要である。また、生成される質問文に関しても、応募者のプライバシーを侵害しない形で質問が具体化されるよう、倫理基準を設定する必要がある。例えば履歴書に記載された情報を過度に詳細に引用することや、応募者が回答を躊躇するような質問文を生成することを避けるためのフィルタリング機能の導入が望ましい。これにより、より安全かつ有用な質問生成システムの設計が可能となる。

6. まとめ

本研究では、情報を引き出すことを目的とした対話場面において効果的な質問文を自動で生成するため、個人の背景情報に基づいた深堀質問生成手法について検討した。評価の結果、提案手法による質問文は一般的な質問文と比較して適切性が高く、より個人へ寄り添うことができ、さらに回答の難易度が低下する傾向にあることが示唆された。

本手法の導入により、面接担当者の負担軽減やインタビューの効率化が期待される。特に、人事面接のような情報収集を目的とした対話において、提案手法が有用であるこ

とが確認された。また、本手法はカウンセリングや窓口相談など、情報取得を目的とする他の対話場面にも応用可能性を有すると考えられる。今後は引用する経歴証拠の精査、被験者数を増加した評価実験等、信頼性の高い質問生成を目指し各種検討を進める。

参考文献

- [1] 土屋 洋輝, 橋本 舜矢, 曾 傑, 坂戸 達陽, 中野 有紀子: “大規模言語モデルを利用した動機づけ面接会話エージェント”, 人工知能学会全国大会論文集, (2024).
- [2] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, et al., “Language Models are Few-Shot Learners”, NeurIPS, (2020).
- [3] Ming-Hsiang Su, Chung-Hsien Wu, Yi Chang, “Follow-Up Question Generation Using Neural Tensor Network-Based Domain Ontology Population in an Interview Coaching System, Interspeech”, (2019).
- [4] 井上 昂治, 原 康平, ララ ディベッシュ, 中村 静, 高梨 克也, 河原 達也, “掘り下げ質問を行う就職面接対話システムの自律型アンドロイドでの実装と評価”, 人工知能学会論文誌, (2020).