

ペルソナ・就活における能力特性を反映した面接対話シナリオ生成

石戸谷由梨 + 小林 一郎 +

† お茶の水女子大学

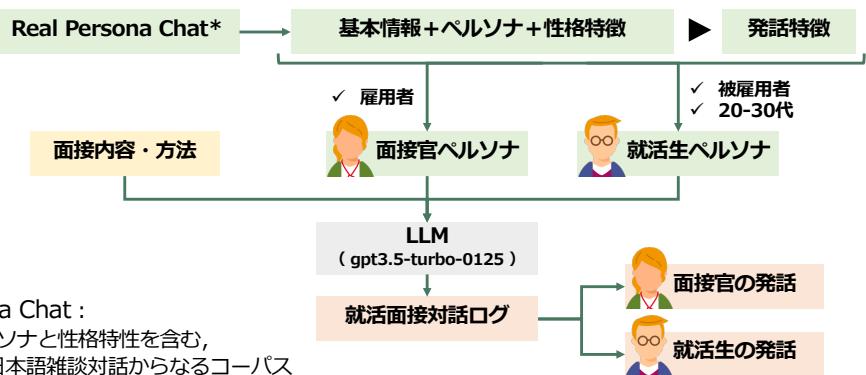
◆ 背景・目的

- LLMの課題
 - バイアスが含まれたアウトプットを生成する危険性
- AI採用の課題
 - バイアス含まれた診断の危険性

多種多様な能力特性を反映した男女の面接対話シナリオ

新たな採用AIの開発

◆ 概要



◆ 対話生成の流れ

プロンプト

- 就職活動を行う学生と面接官が、面接を行っているやり取りを再現してください。
- 面接官は、学生のコミュニケーション能力・主体性・チャレンジ精神・協調性・誠実性を問うような質問を投げかけてください。学生の回答に対して、先の5項目をより明確にするために追加の質問を行うことも可能です。
- 学生は次に示すような人物像を持っている就活生として、質問に答えてください。
※上記の後に文章化したペルソナ情報（右図より作成）を入力

生成した対話

面接官: こんにちは、よろしくお願いします。私は面接官です。まず最初の質問ですが、コミュニケーション能力について教えてください。人のコミュニケーションが得意な部分や工夫していることがあれば教えてください。

就活生: こんにちは、よろしくお願いします。私はコミュニケーション能力に自信があります。具体的には、大学のグループプロジェクトでは積極的に意見を交換し、チーム全体の理解を深めるよう努めています。また、聞き手の気持ちに寄り添ったコミュニケーションを心がけています。

面接官: なるほど、ありがとうございます。次に主体性について質問しますね。どのような状況でも主体的に行動することができる経験や例を挙げただけますか？

就活生: はい、主体性については、大学のサークルでのイベント企画や実行の経験があります。自ら課題を見つけて、解決策を提案し、メンバーと協力して実現することで、私の主体性を示すことができました。

面接官: よくやっているようですね。次はチャレンジ精神についてです。新しいことに挑戦することや困難を乗り越える姿勢がどれくらいあると感じますか？具體例を教えていただけますか？

就活生: チャレンジ精神については、新しい分野の勉強や技術習得に積極的に取り組んでいます。例えば、プログラミングの学習や留学を通じて、自分の可能性を広げるために努力しています。

面接官: 素晴らしい取り組みですね。次に協調性についてお聞きします。チームでの活動や協力の経験を通じて、協調性を発揮した具体例を教えてください。

就活生: 協調性については、大学のボランティア活動やサークルのイベントで、メンバーとの円滑なコミュニケーションや意見調整を行なながら、共通の目標に向けて協力することで、チーム全体の成果に貢献してきました。・・・（以下略）

面接官	就活生
基本情報	基本情報
名前=DD, 性別=女性, 年齢=30-39歳, 教育歴=大学卒, 就労状況=雇用者, 居住地域=東京都	名前=AH, 性別=男性, 年齢=19歳, 教育歴=その他, 就労状況=学生, 居住地域=埼玉
ペルソナ情報	ペルソナ情報
□ 東京都出身です。 □ 韓国語の勉強をしています。 □ 見晴らしの良い場所で、綺麗な景色を眺めるのが好きです。 □ 音楽に合わせて自由に踊るのが好きです。 □ 旅行に行くのが好きです。 □ スピード感のあるゲームが嫌いです。 □ 色んな国の方と友達になりたいです。 □ 公園に行ったり、カフェに行ったりするのが好きです。 □ 世話好きで、人に頼られるでやる気の出るタイプです。 □ 猛烈に素直ではないところが短所です。	□ 私は学生である。 □ 埼玉県出身である。 □ 私は毎日朝食を食べない。 □ 私は毎日ウォーキングをする。 □ 私はよくコンビニに行く。 □ 私はタバコが早い。 □ 自分は物覚えが悪い。 □ 自分は将来の目標が明確に決まっている。 □ ハイセンスなことを見つけられる。 □ 自分は好きなものにはとことんこだわる。
性格特徴	性格特徴
BigFive Openness = 3.67, Conscientiousness = 4.08, Extraversion = 4.00 Agreeableness = 4.17, Neuroticism = 4.83	BigFive Openness = 5.25, Conscientiousness = 3.17, Extraversion = 3.33, Agreeableness = 4.17, Neuroticism = 4.42
KISS18 BasicSkill = 3.67, AdvancedSkill = 2.67, EmotionalManagementSkill = 3.33, OffenceManagementSkill = 3.67, StressManagementSkill = 3.00, PlanningSkill = 3.67	KISS18 BasicSkill = 4.00, AdvancedSkill = 4.33, EmotionalManagementSkill = 4.00, OffenceManagementSkill = 4.00, StressManagementSkill = 4.00, PlanningSkill = 4.67
IOS IOS = 1	IOS IOS = 4
ATQ Fear = 4.14, Frustration = 3.17, Sadness = 5.14, Discomfort = 4.00, ActivationControl = 4.57, AttentionalControl = 4.40, InhibitoryControl = 4.29, Sociability = 5.00, HighIntensityPleasure = 4.66, PositiveAffect = 4.40, NeutralPerceptualSensitivity = 6.00, AffectivePerceptualSensitivity = 5.60, AssociativeSensitivity = 3.60	ATQ Fear = 5.00, Frustration = 3.50, Sadness = 3.00, Discomfort = 3.33, ActivationControl = 3.71, AttentionalControl = 3.80, InhibitoryControl = 3.14, Sociability = 4.00, HighIntensityPleasure = 4.57, PositiveAffect = 3.40, NeutralPerceptualSensitivity = 4.20, AffectivePerceptualSensitivity = 4.40, AssociativeSensitivity = 4.80
SMS Extraversion = 2.40, OtherDirectedness = 2.17, Acting = 4.00	SMS Extraversion = 2.60, OtherDirectedness = 3.58, Acting = 3.75

◆ 実験1：ジェンダーバイアスの追加学習

femalewords list*	女性は 同意的、愛情深い、注意深い であるべき
malewords list*	男性は 活動的、冒険家、攻撃的 であるべき
実施内容 ジェンダーバイアスのある文章を（例：男性は～であるべき）学習させ、故意にジェンダーバイアスを出力に含ませ、バイアスの測定を行う。	
使用データ • Fine-tuning用のデータ : *Reading for Bias (https://github.com/gender-bias/gender-bias/blob/master/README.md) • 分析データ : 事前に作成した男女1人ずつ全4通りの就活面接対話ログ	
学習方法 • gpt3.5-turbo-0125へのLoRA fine-tuning • プロンプトで与える	
バイアス検出 GenBit (https://github.com/microsoft/responsible-ai-toolbox-genbit) • 入力 : テキスト • 出力 : Genbit score, 性別を定義する言葉が含まれる割合 (female gender definition words, male gender definition words, non binary gender definition words, trans gender definition words, cis gender definition words)	

結果

□ 生成した対話

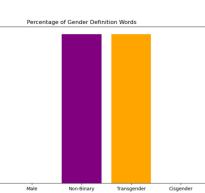
- 対話ログの精度が悪化
- ペルソナの特徴が反映されていない

- 原因① : 短文によるFine-tuning
- 原因② : 毒性を抑えるための出力制限

□ GenBit

- バイアスを検知することができなかった

- 原因 : 性別を定義する特定の言葉のみを感知している



◆ 実験2：構造化面接対話シナリオの生成

実施内容	米国連邦政府人管理局の構造化面接を元に、実験に構造化面接を適応させる。 ※構造化面接 : あらかじめ設定しておいた評価基準・質問項目を基に手順通りに進める面接。反対に、個人に合わせて質問を変更する形式を非構造化面接という。
使用データ	分析データ : 事前に作成した男女1人ずつ全4通りの就活面接対話ログ
実施方法	下記の順番でプロンプトを与え、構造化面接を再現する。 1. 面接ガイドライン（面接官の果たすべき役割、面接官の心構え、採用基準、面接の流れ、面接官がやつけてはいけない行動・態度やNG質問） 2. 面接官と就活生のペルソナ入力 3. 面接で使用する質問と評価基準
結果	□ 出力された対話の内容に、多様性が感じられた □ ペルソナごとに面接評価のばらつきが見られた

◆ 今後の取り組み

□ 実験1

- 出力の精度に影響を与えないFine-tuningの手法の検討
- 精度の高いジェンダーバイアス測定方法の検討

□ 実験2

- 入力と出力の整合性比較
- 非構造化面接と定量的な比較調査