住民属性と認知的対立理論を用いた 生成 AI インタビューによる潜在的な地域情報の収集

佐々 葵[1]　寺島 青[1]　田村 直樹[1]　庄子 和之[2]　浦野 健太[1]　米澤 拓郎[1]　吉川 正[3]　河口 信夫[1,2]

概要：住民が持つ場所の隠れた魅力や場所固有の印象などの地域情報は，観光や地域政策をはじめとする地域振興の多岐にわたる場面で有用であり，その系統的な蓄積は重要な課題である．しかし，主体的な発信を前提とする既存の手法（VGI や SNS 分析等）は，住民が発信しない地域情報を構造的に捕捉できず，かつ収集した情報を蓄積・再利用する仕組みを備えていない．そのため，収集規模を拡大しても質の向上に限界がある．これらの課題に対処するため，本研究では，住民背景に基づく動的質問生成と認知的対立理論（Cognitive Conflict Theory）に基づく深掘りを組み合わせ，地域情報を能動的に引き出す LLM インタビューシステムを提案する．本システムは，住民の持つ地域情報を収集し，参照・再利用可能な形で蓄積し，後続の対話において，より深い地域情報の引き出しに利用するよう設計されている．愛知県新城市の住民 38 名を対象とした実験では，本システムは 74 件の地域情報を収集した．参加者の自己申告では，94%が収集内容を自身ではオンラインに発信した経験がないと回答した．これは，少なくとも当該参加者自身がオンラインで発信していない地域情報を，本システムが対話を通じて引き出しうることを示唆しており，提案手法の実現可能性を初期的に支持する．

1. はじめに

住民が持つ場所の隠れた魅力や場所固有の印象などの地域情報は，観光，地域活性化，定住促進といった地域振興において有用である．これらの地域情報は，観光客にとっては，地域の魅力を知る手がかりとなり，地域政策にとっては，住民のニーズや地域の課題を把握するための重要な情報源となる．複数の地方での大規模な収集・共有活動は，来訪者数の増加や，地方からの人口流出の抑制にも寄与しうる．しかし，これらの地域情報は系統的に収集されておらず，蓄積基盤の構築は重要な課題である．VGI（Volunteered Geographic Information）から SNS データの分析に至るまで，地域情報の収集に関する先行研究はその有用性を示してきた [1], [2]．しかし，これらの手法はいずれも住民の自発的な情報発信を前提としており，「発信できる」ことと「実際に発信している」ことは別である．SNS 利用者のうち実際に発信しているのは約 3 割にとどまる [3]．発信しない住民も，身近な人とは日常会話の中で地域情報を共有している可能性があり，そのような地域情報は既存の手法では構造的に取りこぼされる．本研究ではこのような地域情報を**潜在的地域情報**と定義し，**主観的情報**と**客観的情報**の 2 種に分類する．このうち主観的情報は，観光や定住促進の中心となる地域の体験的魅力を伝えるため特に価値が高く，一方で客観的情報は公開情報源からも入手できることが多い．本研究は主に主観的情報を収集対象とするが，対話の副産物として客観的情報も併せて収集される場合がある．加えて，地域情報はイベントや店舗，地域の慣習が変化するにつれて変容するため，一度の収集ではすぐに陳腐化する．しかし，アンケートや VGI で地域情報を収集しても，それを次回以降の収集のために蓄積し，再利用する基盤は存在しない．こうした基盤がなければ，収集は毎回初期状態から始まり，鮮度を保つことも，規模に応じたカバレッジ拡大もできない．

これらの課題に対処するため，本研究では LLM を用いた対話型インタビューシステムを提案する．図 1 にシステムの概要を示す．潜在的地域情報の収集が依然として困難な理由は，(a) 住民ごとに「何を聞くか」の動的な決定，(b) 表層的回答から個人的体験・主観へと「どう深掘るか」の設計，(c) 得られた情報の後続セッションで再利用可能な形での蓄積の 3 点を同時に満たす必要があるためである．LLM 登場以前のアンケートや人手インタビューは (a) や (c) を欠いていた．LLM 登場後の研究として Cuevas らは，回答者の背景情報を用いた LLM による動的な質問生成自体には一定の有効性を認めつつも，インタビュー全

[1] 名古屋大学大学院 工学研究科
[2] 名古屋大学 未来社会創造機構
[3] トヨタテクニカルディベロップメント株式会社

体を単一の LLM エージェントに委ねた設計では参加者の具体的な動機や個人的経験といった豊かな質的情報を十分に引き出せない場合があると報告している [4]．本研究は，その後の LLM の能力向上に加え，対話を行う LLM に状態遷移を委ねるのではなく状態遷移ロジックを外部の制御モジュール（Dify ワークフロー上のインタビュー制御モジュール）に切り出す構成を採り，この制約の緩和を目指す．そのうえで，3 要件に対応する 3 要素を統合する．(a) **プロフィール適応型質問生成**では，住民の背景プロフィールから精通するトピックを推定し，それに応じて質問を生成し，住民ごとに動的な質問生成を可能にする．(b) **認知的対立理論（Cognitive Conflict Theory）に基づく深掘り**は，他の住民の意見との対峙を提示し，参加者がそれまで意識していなかった主観の言語化を促す [5]．(c) **循環的蓄積機構**は，得られた情報をデータベースに記録し後続セッションでの参照を可能にする．3 要素は単独で機能するのではなく，循環的蓄積機構が認知的対立の刺激源となる情報を供給し，プロフィール適応型質問生成が蓄積データから関連情報を引き出し，認知的対立に基づく深掘りが蓄積データを濃くするという相互補完関係に新規性がある．

愛知県新城市の住民 38 名を対象とした実験では，本システムは 74 件の地域情報（LLM により区切られたもの）を収集した．セッション後アンケートにおける参加者の自己申告では，94%が収集内容を自身ではオンラインに発信した経験がないと回答し，72%は家族や近隣住民と口頭で共有していた．これらの結果は，少なくとも当該参加者自身がオンラインに発信していない地域情報を，プロフィール適応型質問生成と認知的対立理論に基づく深掘りを組み合わせた対話を通じて，提案システムが引き出しうることを示唆する．

本論文の貢献は以下の通りである．

- オンライン発信に依存する既存手法では未発信住民の地域情報を構造的に捕捉できない課題を整理し，これを潜在的地域情報として概念化した．
- プロフィール適応型質問生成・認知的対立理論に基づく深掘り・循環的蓄積機構を統合した LLM インタビューシステムを設計・実装した．
- 提案手法による地域情報収集実験を実施し，当該参加者自身がオンラインに発信したのない地域情報を本システムが引き出しうることを示唆した．

2. 関連研究

2.1 地域情報収集手法

地域情報の収集手法はこれまで大きく変化してきた．Goodchild が提唱した VGI は，市民の自発的な情報提供によって低コストでの大規模収集を実現し [1]，近年では SNS や口コミサイトの投稿を分析して主観的な地域情報を取り出す研究も数多く報告されている．例えば，TripAdvisor

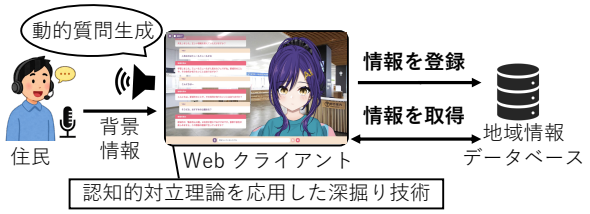


図 1 提案 LLM インタビューシステムの概要．住民の背景情報を起点に，Web クライアント上で（a）動的質問生成，（b）認知的対立理論に基づく深掘り，（c）地域情報データベースへの登録・再取得（循環的蓄積）の 3 要素が連動する

レビューに対するトピック抽出による観光体験を構成するテーマの特定 [2]，オンラインレビューからの観光地イメージ（destination image）の多面的分析 [6]，アスペクトベース感情分析による訪問者の不満要因の抽出 [7]，さらにジオタグ付き写真の顔表情を画像認識により解析した場所単位の感情マッピング [8] といったアプローチが提案されており，従来は定量化が難しかった場所への主観的評価やアタッチメントの定量的記述に貢献している．しかし，これらはいずれも住民の自発的な情報発信を前提としている点で共通する．そのため Zheng らが「我々が得られるのは常に一部のユーザからのサンプルデータであり，すべてのユーザがデータを共有するわけではない」と指摘するように [9]，プラットフォーム側は誰が投稿するかを制御できず，得られる地域情報は発信意欲の高いユーザに偏らざるを得ない．これは VGI 登場当初から指摘されてきた根本的な課題であり，SNS 時代になっても解消されておらず，発信しない住民の地域情報は依然として既存手法に構造的に捕捉されていない．そこで提案システムは対話を通じて情報共有を促し，インターネットでの発信意欲が薄い住民から地域情報を引き出す．これにより，地域情報収集の技術的な障壁を低減する．

一方，地域情報は時間とともに変化するため，一度の収集では鮮度を保てない．Niedomysl [10] および Kitagawa [11] は，地域の魅力を潜在的住民に対して長期的に発信し続けることが重要であると示しており，地域情報の継続的な更新と蓄積が不可欠である．しかし，既存の手法には収集した地域情報を後続のセッションで蓄積・再利用する仕組みがなく，収集は毎回初期状態から始まる．提案システムは，これら 2 つの課題に対処するために循環的蓄積機構を備える．すなわち，地域情報をリアルタイムで蓄積し，後続のセッションでの再利用により，収集データの鮮度と累積性を維持するよう設計されている．

2.2 人間から情報を引き出す手法

対面のインタビューでは，回答者が「他者からどう見られるか」を意識して否定的評価につながり得る回答を控えたり，社会的に望ましいとされる回答に寄せてしまう社会

的望ましさバイアスが生じやすい．地域情報の収集においても，住民が地元への批判的評価や個人的な不満を控え，当たり障りのない肯定的情報のみを語る傾向が想定される．Lucas らは，バーチャルヒューマンという人間がロボットとして振る舞う対話実験において，参加者が対話相手を人間操作ではなくコンピュータ制御であると信じた条件のほうが，自己開示への恐れと印象管理（悪く見られないよう自身を装う傾向）が低下し，より率直な発話と感情表出が観察されることを示した [12]．提案システムは，この知見に基づき LLM による対話を収集手段として採用する．

固定質問による調査は，住民ごとの知識領域の違いに対応できず，表面的な回答にとどまる．Cuevas らは，回答者の背景情報を用いた LLM による動的な質問生成が，表面的な回答を抑制し豊かなデータを引き出す上で一定の有効性を持つことを示した [4]．ただし同研究は，インタビュー全体を LLM に委ねた単一エージェント型の設計では引き出せる情報の豊かさに限界があることも報告しており，本研究ではこの限界への対処として 1 章で述べた状態遷移ロジックの外部化を採用している．提案システムは，背景情報に基づく動的質問生成という有効な側面を住民のプロフィールからのトピック推定に特化させたプロフィール適応型質問生成として継承し，より詳細な地域情報の収集を可能とする．

主観的情報は住民自身が無意識に保持している場合が多く，単純な問いかけでは引き出しづらい．認知的対立理論 [5] は，異なる見解との対峙が自己の認識への内省を促し，自身の知識のより深い言語化を引き出すことを示している．Yao らはこれを LLM の文化的アライメント用データ構築に応用し，異文化間の差異を用いて自文化への内省を促している [13]．また，Jo らは過去の対話の参照が同一ユーザ内における長期的な自己開示の深化を促すことを示した [14] が，その記憶利用は個々のユーザ自身の過去対話に限定されていた．対立を生起させる方法には様々な設計が考えられ，住民によっては表現や提示の仕方により深掘りの効果が異なる可能性もある．そのため，本研究ではまずその設計空間のうち最も基本的な形式，すなわち他の住民の意見をそのまま提示する形式を採用し，その有効性の初期的な検証を目指す．具体的には循環的蓄積機構により収集された他住民の地域情報を対立意見として再利用し，Jo らが示した記憶活用を異なるユーザ間に拡張するように設計する．他の設計の方向性については 5.3 節で議論する．

3. 提案手法

3.1 システム概要

提案システムは，地域住民と音声対話を通じてインタビューを行い，潜在的地域情報を収集する自動化されたパイプラインである．本システムは主観的情報の収集を主目的として設計されており，認知的対立理論を用いて個々の

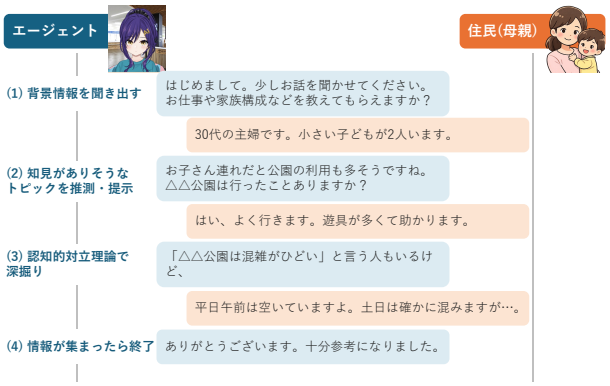


図 2 提案システムによるインタビューフローの例．プロフィール収集・トピック決定・認知的対立理論に基づく深掘り・終了の 4 ステップで構成される．図中の対話内容は説明用に簡略化した例示であり，実際の応答は LLM がセッションごとに動的に生成する．また図では各ステップを 1 ターンで描いているが，実セッションでは複数ターンにわたる場合がある．

意見の深掘りを行う．本システムは以下の 3 つの設計原則に基づく．

(1) プロフィール適応型質問生成：住民のプロフィールから精通するトピックを推定し，それに応じた質問を生成する
(2) 認知的対立理論に基づく深掘り：他の住民の意見を提示し，より詳細かつ具体的な回答を引き出す
(3) 循環的蓄積：収集した地域情報を即座にデータベースに登録し，後続のセッションで再利用する

3.2 インタビューフロー設計

インタビューは以下の 4 ステップで構成される．具体的な会話内容のイメージを図 2 に示す．

(1) **プロフィール収集**：セッション開始時に，年齢・性別・家族構成・居住年数の基本属性を会話形式で収集する．提供したくない属性がある場合は「答えたくない」と回答できるようにする．
(2) **トピック決定**：プロフィールに基づき，LLM が住民の精通するトピックを推定し，対話の中で提示する．例えば，子育て中の親には公園の遊具情報，会社員には職場近辺の飲食店情報を提示する．住民が地域情報を持つトピックが特定された段階で，次のステップに進む．
(3) **認知的対立理論に基づく深掘り**：本システムは 5.3 節で議論する対立提示のデザインスペースのうち，最も基本的な形式を採用し [5]，住民の主観的情報および当該トピックに関する個人的体験を引き出す深掘りを行う．具体的には，LLM が類似した背景を持つ他の住民の同じトピックに関する意見を検索し，以下のように提示する：「〇〇って言っている人もいるけど，あなたはどう思う？」これにより，住民は提示された意見と自身の経験を比較し，より具体的かつ個人的な情

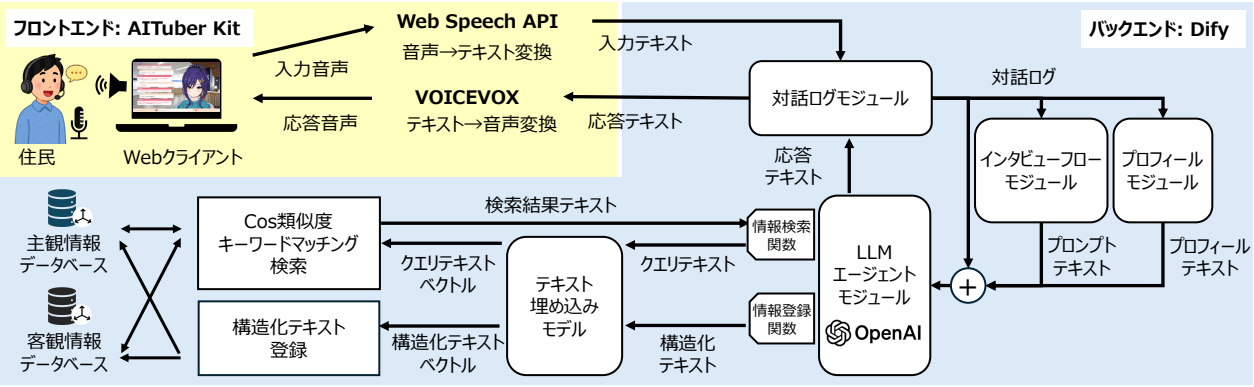


図 3 システムアーキテクチャ：フロントエンド（Web Speech API・VOICEVOX・Dify チャットフロー）とバックエンドモジュール（フロー制御・検索・地域情報記録）

報を共有するよう促される．

(4) **終了**：閾値以上の地域情報が収集できるまで深掘りを継続し，特定の時点でセッションを終了する．実験中は，セッション時間の制約に応じてこの閾値を適宜 2 件から 3 件の間で調整した．特定の時点で地域情報が獲得されたかどうかは，セッション中に LLM がリアルタイムで判定し，地域情報の区切りも LLM が判定する．獲得された地域情報はリアルタイムでデータベースに保存され，後続のセッションで再利用される．

3.3 技術アーキテクチャ

システムアーキテクチャを図 3 に示す．本システムはユーザインタラクションを処理するフロントエンドと，インタビュープロセスを管理し地域情報を蓄積するバックエンドから構成される．

3.3.1 フロントエンド

フロントエンドは AITuber Kit [15] をベースとした Web クライアントであり，著者らが軽微な改修を加えている．アバターキャラクターが対話を進行し，チャットログパネルに対話履歴を表示する．さらに以下の 2 つのコンポーネントと連携し，音声入力・音声応答による対話システムを実装する．1 つは Web Speech API による音声認識，もう 1 つは VOICEVOX [16] による音声合成である．

3.3.2 バックエンド

バックエンドは LLM を用いたシステム開発基盤である Dify [17] を用いて，インタビューセッションの管理（Dify チャットフロー）と地域情報の蓄積（Dify ナレッジベース）を行うモジュールを実装した．対話ログモジュールは対話履歴をリアルタイムで保存し，必要に応じて他のモジュールに提供する．インタビュー制御モジュールは LLM に送信するプロンプトを切り替えることでインタビューの流れを制御し，そのフローは前述のインタビューフロー設計に従う．プロフィールモジュールはセッション中に収集した住民のプロフィールを保持し，プロンプトに付加する．LLM エージェントモジュールは，現在のユーザ発話を含む対話ログ，プロフィール文，現在のインタビューステップ用のプロンプトを入力として応答を生成する．LLM エージェントモジュールは必要に応じて情報検索関数および情報記録関数を呼び出す．LLM は Dify エージェントノードにより自律的に関数を呼び出し，関数は Dify ワークフローを用いて実装されている．情報検索関数は LLM が生成した検索クエリをベクトル化し，コサイン類似度およびキーワードマッチングを用いてデータベースから関連地域情報を検索する．情報記録関数は LLM が生成した新規地域情報を「○○な人は△△に関して□□と感じている」という形式に整形してベクトル化し，その地域情報を提供した人物のプロフィールとともにデータベースに登録する．この時の○○な人は，例えば「在住 20 年，40 代子連れの男性」（年齢・性別・家族構成・居住年数）などのインタビューの中で収集された属性を組み合わせたものである．

データベースには Dify のナレッジベースを用いる．データベースは情報記録関数によって更新され，情報検索関数によって検索される．設計思想としては，主観的情報と客観的情報は，必要となる場面や，情報の信頼度等の特性が異なるため，分類し別々での蓄積が望ましいと考えられる．ただし現状は自動分類モジュールが未実装であり，収集した地域情報はすべて話者属性とともに単一のデータベースに蓄積される．自動分類は今後の課題である．

4. 実験と結果

本節では，提案システムによる潜在的地域情報の引き出しの実現可能性を検証するための実験を実施した．愛知県新城市の住民を対象に，提案システムによる音声対話インタビューによる実験と，比較対象として人手によるワークショップ形式の地域情報収集実験を行った．両実験で収集されたデータは，収集事例の定性的観察（4.2 節），知識グラフによる定量比較（4.3 節），参加者のアンケート（4.4 節）の 3 観点から分析する．以下，各実験の設定を順に述べる．



図 4 実験の様子．（上）ベースラインの人手による収集：参加者は付箋で地域情報を記入し，新城市の地図および分類済み掲示板に集約する．（下）提案システムによる収集：住民はヘッドセット付きノート PC 上で音声対話インタビューを行う．いずれも新城市西部公民館にて実施

4.1 人手による収集（ベースライン）

提案システムを用いずに人手で地域情報を収集する事前実験を実施した（2025 年 11 月，新城市西部公民館にて開催されたワークショップ）．実験の様子を図 4 の上段に示す．収集手順は以下の 3 ステップである．(1) 住民がおすすめのスポットとその特徴を付箋に書き，掲示板に貼る．(2) スタッフ 1 名が時間に余裕のある場合に限り，その付箋を見て対話を通じて深掘りを行う．(3) 深掘りで得られた地域情報をスタッフが付箋に書き起こし，掲示板に貼る．ワークショップ全体の収集時間は約 4 時間，参加者は推定 30〜40 名で，1 人あたりの対話時間は 0〜5 分であった．最終的に約 70 件の地域情報が収集された．収集された付箋に対して，4.3 節で説明する同一の Atomic Fact 分解およびトリプル抽出パイプラインを適用して知識グラフを構築した．本実験は，スタッフが常時必要であり参加者全員に対して一様な深掘りを行えない点で，継続的・大規模な収集には不向きであるという課題を内在している．

4.2 提案システムによる収集と結果

愛知県新城市において，住民 38 名を対象とした提案手法による地域情報収集実験を実施した（2026 年 2 月，新城市西部公民館）．実験の様子を図 4 の下段に示す．すべての参加者には研究・開発目的・データの利用方針・匿名化処理を説明し，事前に同意を得て実施した．フロントエンドを 2 台並列で稼働させ，1 人あたりのセッション時間は 5〜10 分とした．なお，バックエンドは両フロントエンドで共有される構成である．セッション後アンケートには 36 名から回答を得た．実験全体の収集期間は約 3.5 時間であった．

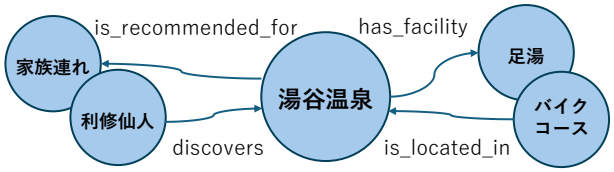


図 5 抽出されたトリプルの例：「湯谷温泉」を中心エンティティとした関係の一部．円がエンティティ，矢印付きの辺がリレーション（述語）に対応する

実験では 74 件の地域情報が収集された（情報の区切りは LLM が判定）．収集データには個人的な体験談や具体的なエピソードが含まれた．例えば次のようなものである．「新城総合公園は，40 代子連れ男性の観点で見ると，遊具が非常に充実しており，長い滑り台が子どもたちに大人気．ピクニックを兼ねてお弁当を持参しても楽しめ，知らない子どもと自然に順番を守るなど社会性も学べる．一方，夏場は非常に暑くなるため，飲み物の持参が必須．」

著者らが各地域情報から複数のキーワードを抽出し，それらを OR 結合した検索クエリで予備的なインターネット検索を実施したところ，一部の地域情報はインターネット上で確認できなかった．例えば次のような事例である．「桜淵は桜の名所であるが，老木が年々目立つようになるため，若木を順次植えて景観を維持してほしい．」「新城納涼花火大会は毎年コンセプトが変わるため飽きない．」「市川の桜の名所はその美しさにおいて，吉野山の千本桜の小規模版のようである．」

4.3 情報量の比較

この節では，提案手法で収集された地域情報の情報量を定量的に分析する．比較のベースラインとして，提案システムを使用せず人手で実施したワークショップ形式の収集記録を用いる．分析手法には，LLM を用いた動的時系列知識グラフ（TKG）構築手法である ATOM [18] を参考にした．ATOM は入力文書をまず 1 事実 1 単位の最小自己完結スニペット（Atomic Fact）に分解し，各 Atomic Fact から (主語, 関係, 目的語, 開始時刻, 終了時刻) の 5-tuple を並列抽出する手法である．Atomic Fact への分解は，長文から直接構造化抽出を行う際に LLM が陥る「忘却効果 (forgetting effect)」，すなわち事実の取りこぼしを回避する効果が報告されている．本研究の現行システムは時系列情報を含めた構造化を行っていないため，本節の比較では ATOM のうち Atomic Fact 分解の枠組みのみを流用し，各 Atomic Fact から (主語, 述語, 目的語) の 3 つ組 (トリプル) を抽出し，知識グラフを構築した．トリプルの要素は，コサイン類似度をもとに同一エンティティを結合し，知識グラフのノード（エンティティ）とエッジ（リレーション）を形成する．提案システムで収集した地域情報および人手によるワークショップでの収集記録の双方に同一

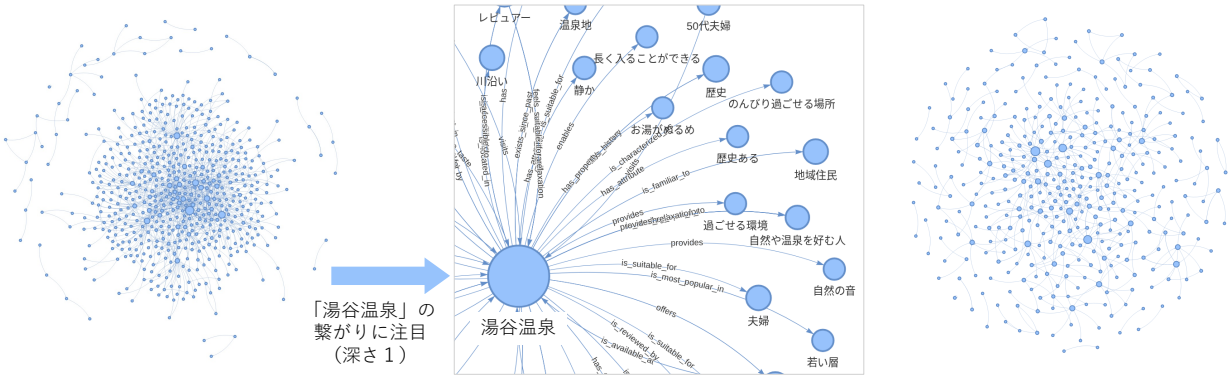


図 6　提案手法（左図）と人手による収集（右図）から構築した知識グラフの全体像．中央図は左図のうち「湯谷温泉」エンティティを中心とした周辺関係の拡大表示であり，1 つの対象に複数のリレーションが付与され多面的に記述されている様子を示す

表 1　提案システムおよびベースライン（人手による収集）から構築した知識グラフの統計（参考値；両条件は実施日・対話時間が異なるため厳密な統制比較ではない）

手法	エンティティ数	リレーション数	平均関係数
提案システム	583	1,188	2.04
人手による収集	381	427	1.12

プロンプトを適用し，両手法で得られた情報の総量をエンティティ数・リレーション数・1 エンティティあたりの平均リレーション数である平均関係数の 3 指標で比較する．プロンプトは ATOM で公開されているプロンプトをもとに，Atomic Fact 分解時は冒頭に日本語で出力する旨を追加したものを用い，トリプル抽出時のリレーションは原版に倣い英語（is_recommended_for や has_facility 等）で出力するようにした．これは，プロンプト内で列挙されているリレーション例を改変せずに使用し，LLM にその制約を忠実に再現させるためである．

抽出されたトリプルとそれを用いた知識グラフの構成を具体的に示す．図 5 は，本実験で実際に得られたトリプルの一部を抜粋し，1 つの中心エンティティ（湯谷温泉）の周辺関係を可視化した例である．ここでは「湯谷温泉が家族連れ向けである（is_recommended_for）」「足湯を備える（has_facility）」「利修仙人によって発見された（discovers）」等の関係が示されている．このように，各トリプルは 2 つのエンティティを 1 本のリレーションで結ぶ最小単位を成し，これらの共通エンティティを介する連結により知識グラフが形成される．構築した提案手法・ベースライン全体の知識グラフを図 6 に示し，各指標を表 1 に示す．

表 1 より，提案システムはエンティティ数で約 1.5 倍，リレーション数で約 2.8 倍の情報を収集できたことが分かる．平均関係数も，提案システムが 2.04 であるのに対し，人手による収集は 1.12 であった．特にリレーション数の差はエンティティ数の差より大きく，平均関係数も提案システムの方が高い．これは，提案システムが単に言及対象（場所・物・人）を多く引き出すだけでなく，それぞれの

表 2　収集地域情報の潜在性に関するアンケート結果（$n = 36$）

質問	「ある」	「ない」
家族や地域の人に話した経験があるか？	72%	28%
ネット等の公の場で発信した経験があるか？	6%	94%

対象について複数の特徴や関係性を深掘りできていることを示唆する．ベースラインでは発話が「場所名+短い特徴」にとどまる事例が多かったのに対し，提案システムでは認知的対立理論による深掘りを通じて，同一の対象に対する利用シーンや主観的評価，関連エピソードといった多面的な情報が収集された結果と考えられる．ただし，両者は参加者・実施日・収集形式が異なる実験の比較であり，厳密な統制比較ではない点は留意が必要である．

4.4　収集情報の潜在性に関するアンケート結果

参加者に共有行動についてアンケートを実施した（表 2）．72%が地元の会話の中で内容を共有した経験があり，94%は自身ではオンラインに未発信であると回答した．これは，本システムが住民が会話の中では共有するがオンラインでは積極的に発信しない地域情報を収集していることを示唆する．ただし，この結果はあくまで各参加者自身の発信行動を反映するものである．本システムにおける最終的な目標は，収集された内容が公開されているオンライン情報源から発見されないことである．著者らによる予備的なインターネットキーワード検索では一部の内容を検出できず，この目標を裏付ける定性的な証拠を得た．信頼性のある潜在性指標の確立は今後の課題である．本結果は自己申告に基づくため，記憶バイアスや社会的望ましさの影響は否定できない．

4.5　認知的対立理論に基づく深掘りの動作例

この節では，認知的対立理論に基づく深掘り設計の動作例を示す．表 3 に，認知的対立理論に基づく深掘り設計の

表 3 認知的対立理論に基づく深掘り設計の動作例（五平餅茶屋に関する会話抜粋）

発話者	内容
住民	五平餅は味噌だけのが好きで，くるみとか入っているのは好きじゃない．
LLM（他者意見提示）	味噌だけのシンプルなのが好きなんですね．ちなみに，「五平餅は味が濃い」って言っている人もいるんですが，濃さはどう感じましたか？ちょうど良かったですか，それともあっさり目が好みですか？
住民（深化後）	ちょうど良かったです．時々，家でも作ります．

動作例として，ある五平餅茶屋に関する会話の抜粋を示す．LLM が他者の意見を提示する前は，住民は短い嗜好の表明にとどまっていた．他者意見を提示した後の応答は，個人的な評価および新たな地域情報（自宅で作るという情報）を含む発話に拡張された．これは意図した機構が実際に動作していることを示す一例である．

認知的対立理論に基づく深掘りが，発話内容の深化にどれほど寄与しているかの定量的な評価は重要である．例えば，発話の長さや情報の新規性などの指標での効果の定量的な評価は今後の課題である．

5. 議論

5.1 本研究の制約

本研究は探索的な予備実験であり，いくつかの制約を抱えている．第一に，実験は愛知県新城市の住民 38 名を対象とした単一地域・単一日程での実施であり，地域特性や参加者層の偏りが結果に与えた影響については検証できていない．第二に，潜在性に関する根拠は参加者の自己申告アンケートに依拠しており，記憶バイアスや社会的望ましさバイアスの影響を否定できない．収集された地域情報が公開情報源のいずれにも存在しないという厳密な意味での潜在性は，本研究では検証できていない．第三に，ベースラインとして用いた人手によるワークショップ形式の収集は，提案システム実験とは実施日・参加者・形式が異なり厳密な統制比較ではない．本論文での比較はワークショップ全体での収集効率の参考値として読まれるべきである．

5.2 循環構造の検証限界

本システムの中核的な要素である循環構造，すなわち収集された地域情報の即時蓄積と再利用の効果の検証は重要である．しかし，今回の実験は規模が限定的であり，情報の循環の十分な確認とその効果の検証はできなかった．この結果については，セッション間の会話内容が多様で内容の重複が少なく，参加者の属性も多様で，セッション数が少なかった等の要因が考えられる．しかし，提案システムの設計上，地域情報は収集を重ねるたびに蓄積していくため，次回の実験では循環構造を十分に観測し，効果を検証できる可能性がある．

5.3 認知的対立提示のデザインスペース

2.2 節で述べたように，認知的対立の生起のさせ方には複数の設計が存在し得る．他者の意見を提示する形式の中でも，(i) 提示する意見の選択基準（類似プロフィール vs 異なるプロフィール），(ii) 提示の強度（中立的紹介 vs 明示的な反論型），(iii) 提示する意見の数（単一意見 vs 複数意見の対比），(iv) 出典の文脈化（匿名 vs 属性付与），など多様な操作化が考えられる．本研究はこれらのうち最も基本的な構成（類似プロフィール，中立的紹介，単一意見，属性付与）の有効性を初期的に検証した．住民属性や対象トピックによって有効な提示方法は異なる可能性があり，今後はこれらの設計方針の体系的な探索を計画している．

さらに，これらの設計方針はテキストによる対立提示を前提としているが，対立の生起にはマルチモーダルな手段も検討可能である．例えば，観光案内サイトやイベントポスターなどの視覚的コンテンツを提示しながらそれに対する他の住民の反応や意見を併せた表示が，言語のみでは得られない視覚的文脈を伴う対立を生起できる可能性がある．これにより住民は対象の場所やイベントを具体的にイメージしながら自身の意見を述べやすくなり，より具体的かつ感情を伴う発話の引き出しが期待できる．

また，他者の意見を介さず，過去の地域状況の提示や反実仮想（「もし○○がなかったら地域はどう変わっていたか」）等を介して，自己の既存スキーマと提示内容との間の対立を生起させることも考えられる．このアプローチは他者からの評価という社会的圧力を伴わないため，社会的望ましさバイアスを抑制しつつ深い言語化を促せる可能性があり，今後の重要な探索方向の一つである．

5.4 技術的拡張の方向性

設計面では，NLP 技術を用いて発話を主観・客観に自動分類する仕組み，および蓄積知識ベースとの類似度に基づく適応的な対話終了判定を実装する余地がある．加えて，知識ベース表現の拡張を計画している．現状の提案システムは収集情報をテキスト埋め込みベクトルとして単一データベースに蓄積する設計で，時系列情報や地域情報同士の関係性を明示的に表現しない．地域情報は店舗の閉店・イベントの変更・地域慣習の変容といった形で時間とともに移り変わるため，情報の更新等の時系列性を扱える表現が望ましい．また，地域情報同士の関係性の明示的な表現により，地域情報の構造化や推論が可能になると考えられる．そのため，今後は 4.3 節で参考にした ATOM [18] の 5-tuple TKG 表現を知識ベースに導入し，各事実の妥当性期間 $(t_{\text{start}}, t_{\text{end}})$ を保持して，古くなった情報の検出と更

新の自動化を目指す．また，評価面では，アブレーションスタディ（認知的対立理論のあり／なし）により，発話長や情報の新規性などの指標で認知的対立理論の有効性の定量的な評価を行う．また，信頼性のある潜在性指標を構築し，本システムが真に潜在的な地域情報を引き出すことに成功した事例の分析を進める．

5.5 社会実装と倫理的考慮

収集した地域情報を活用するため，住民や自治体に価値を還元するシステムとの統合を進める必要がある．蓄積してきた地域情報を，1. ライフスタイルの変化などで情報を必要とする住民に対して，2. 政策立案に役立てるために自治体に対して，3. 地域の魅力を伝えるために来訪者に対して共有し，地域発展に貢献するシステムを開発する．この時，社会実装する上では倫理的な課題も考慮する必要がある．収集された地域情報の中には個人を特定できる情報やセンシティブな内容が含まれる可能性があり，これらを適切に匿名化・フィルタリングする仕組みを構築し，プライバシー保護を確保する必要がある．また，悪意を持ったユーザが虚偽の情報を提供するリスクもあるため，収集された地域情報の信頼性を評価する方法も検討する必要がある．さらに，収集された地域情報をどのように共有するかについても，全ての情報を公にすることは望ましくなく，住民の同意や意図を汲み取るプロセスを設計する必要がある．

6. 結論

本研究では，プロフィール適応型質問生成・認知的対立理論に基づく深掘り・循環的蓄積機構の 3 要素を統合した LLM インタビューシステムを提案した．認知的対立提示については，設計空間のうち最も基本的な形式（他の住民の意見の単純提示）を採用し，その有効性の初期検討を行った．住民 38 名を対象とした愛知県新城市の実証実験を実施し，74 件の地域情報を収集した．参加者の自己申告では 94%が収集内容を自身ではオンラインに発信した経験がないと回答した．この結果は，本システムが少なくとも当該参加者自身がオンラインで発信していない地域情報を対話を通じて引き出しうることを示唆し，提案手法の実現可能性を初期的に支持する．今後は，認知的対立提示の他の設計の探索，本研究で議論した制約への対応，循環構造の有効性の実証を進める．さらに，収集された地域情報を住民や自治体，来訪者と共有するシステムの開発を通じて地域社会への還元を進める．

謝辞 本研究の一部は内閣府 SIP3 (JPJ012495)，JST CREST (JPMJCR22M4)，JST RISTEX (JPMJRS23K) に支援いただいています．

参考文献

[1] Goodchild, M. F.: Citizens as sensors: the world of volunteered geography, *GeoJournal*, Vol. 69, No. 4, pp. 211–221 (2007).

[2] Taecharungroj, V. and Mathayomchan, B.: Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand, *Tourism Management*, Vol. 75, pp. 550–568 (2019).

[3] 株式会社 NTT ドコモ モバイル社会研究所：モバイル社会白書 2025 年版, 資料 2-6, , 入手先 ⟨https://www.mobaken.jp/whitepaper/wp25.html⟩ （参照 2026-04-29）.

[4] Cuevas, A., Scurrell, J. V., Brown, E. M., Entenmann, J. and Daepp, M. I. G.: Collecting Qualitative Data at Scale with Large Language Models: A Case Study, arXiv (2024).

[5] Cosier, R. A. and Rose, G. L.: Cognitive conflict and goal conflict effects on task performance, *Organizational Behavior and Human Performance*, Vol. 19, No. 2, pp. 378–391 (1977).

[6] Marine-Roig, E.: Destination Image Analytics Through Traveller-Generated Content, *Sustainability*, Vol. 11, No. 12 (2019).

[7] Ali, T., Marc, B., Omar, B., Soulaimane, K. and Larbi, S.: Exploring destination's negative e-reputation using aspect based sentiment analysis approach: Case of Marrakech destination on TripAdvisor, *Tourism Management Perspectives*, Vol. 40, p. 100892 (2021).

[8] Kang, Y., Jia, Q., Gao, S., Zeng, X., Wang, Y., Angsuesser, S., Liu, Y., Ye, X. and Fei, T.: Extracting human emotions at different places based on facial expressions and spatial clustering analysis, *Transactions in GIS*, Vol. 23, No. 3, pp. 450–480 (2019).

[9] Zheng, Y., Capra, L., Wolfson, O. and Yang, H.: Urban Computing: Concepts, Methodologies, and Applications, *ACM Trans. Intell. Syst. Technol.*, Vol. 5, No. 3 (2014).

[10] Niedomysl, T.: TOWARDS A CONCEPTUAL FRAMEWORK OF PLACE ATTRACTIVENESS: A MIGRATION PERSPECTIVE, *Geografiska Annaler: Series B, Human Geography*, Vol. 92, No. 1, pp. 97–109 (2010).

[11] Kitagawa, F.: Lessons from the Japanese 'regional revitalisation' – impacts of regional industrial policies through decentralisation, *Contemporary Social Science*, Vol. 19, No. 4, pp. 602–621 (2024).

[12] Lucas, G. M., Gratch, J., King, A. and Morency, L.-P.: It's only a computer: Virtual humans increase willingness to disclose, *Computers in Human Behavior*, Vol. 37, pp. 94–100 (2014).

[13] Yao, J., Yi, X., Wang, J., Dou, Z. and Xie, X.: CAReDiO: Cultural Alignment via Representativeness and Distinctiveness Guided Data Optimization, arXiv (2026).

[14] Jo, E., Jeong, Y., Park, S., Epstein, D. A. and Kim, Y.-H.: Understanding the Impact of Long-Term Memory on Self-Disclosure with Large Language Model-Driven Chatbots for Public Health Intervention, *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, Association for Computing Machinery (2024).

[15] tegnike: aituber-kit, GitHub repository (online), available from ⟨https://github.com/tegnike/aituber-kit⟩ (accessed 2026-04-14).

[16] VOICEVOX: voicevox, GitHub repository (online), available from

⟨https://github.com/VOICEVOX/voicevox⟩ (accessed 2026-04-14).

[17] langgenius: dify, GitHub repository (online), available from ⟨https://github.com/langgenius/dify⟩ (accessed 2026-04-14).

[18] Lairgi, Y., Moncla, L., Benabdeslem, K., Cazabet, R. and Cléau, P.: ATOM: AdapTive and OptiMized dynamic temporal knowledge graph construction using LLMs, *Findings of the Association for Computational Linguistics: EACL 2026* (Demberg, V., Inui, K. and Marquez, L., eds.), Rabat, Morocco, Association for Computational Linguistics, pp. 950–966 (2026).