

大規模言語モデルを用いた旅行体験の可視化と代替経路の生成

Visualizing Travel Experiences and Generating Alternative Routes Using Large Language Models

○遠藤 純音 (NAIST) *¹ 久田 翔平 (NAIST) *² 若宮 翔子 (NAIST) *³ 荒牧 英治 (NAIST) *⁴

*¹ Sunao Endo, Nara Institute of Science and Technology, 8916-5, Takayama-cho, Ikoma, Nara 630-0192, endo.sunao.et6@naist.ac.jp

*² Shohei Hisada, Nara Institute of Science and Technology, 8916-5, Takayama-cho, Ikoma, Nara 630-0192, s-hisada@is.naist.jp

*³ Shoko Wakamiya, Nara Institute of Science and Technology, 8916-5, Takayama-cho, Ikoma, Nara 630-0192, wakamiya@is.naist.jp

*⁴ Eiji Aramaki, Nara Institute of Science and Technology, 8916-5, Takayama-cho, Ikoma, Nara 630-0192, aramaki@is.naist.jp

キーワード: 大規模言語モデル (LLM), 旅行記, 経路推薦, 感情分析, 自然言語処理

1. はじめに

近年, Web 上には個人の旅行体験を記録したテキストデータが大量に蓄積されており, これらは観光分野における新たなインサイトの源泉として期待されている. これらのデータを活用し, 旅行体験を分析する先行研究は存在するが, 多くは「訪問場所の可視化 (経路分析)」と「体験内容の分析 (感情分析)」を独立して扱ってきた. 例えば, 従来の経路分析研究⁽¹⁾は, 訪問した地点の地理的關係や多くの人々の移動の全体的な傾向を示すに留まり, その場所で旅行者が具体的にどのような感情を抱いたかという主観的な体験の質を同時に示すことは困難であった. また, 従来の感情分析研究⁽²⁾⁽³⁾は, 訪問場所ごと, あるいは旅行記全体のポジティブ・ネガティブ判定やトピック別の満足度分析が主であり, 「旅行行程のどの場所で」「なぜ満足度が高まったかあるいは低下したか」を時系列に特定することは困難であった.

本研究の目的は2つある. 第一の目的は, 旅行体験を「場所」と「行動 (体験)」と「感情 (体験の質)」の三要素を関連付けた時系列データとして構造化し, 可視化することである. これにより, 旅行者が特に高い満足度を得たポジティブな体験の要因や, 反対に満足度を低下させるネガティブな体験 (本研究で用いるデータセットに含まれる旅行体験の記述はほとんどがポジティブなものであり, ネガティブな体験は少ないため, ここでは旅行全体のうち満足度を低下させている一部分という意味で**満足度のボトルネック**とよぶ. したがって, 仮に旅行記の開始点や終端のみがネガティブな体験であっても, ボトルネックとよぶ) が発生した具体的な場所や要因を, 客観的に特定することが可能となる. しかし, 単に過去のネガティブな体験を特定するだけでは, 利用者が未来の旅行を計画する際の直接的な支援には繋がらない. 従来の観光推薦システムは, 利用者の嗜好プロファイルに基づき, 「行くべき場所 (点)」を推薦するものが主である⁽⁴⁾. 一方で, 旅行中に満足度が低下した区間やネガティブな要素を含む経路に対して, 「もしあの時, 別の行動をしていたら (What-if)」という仮想的な文脈に基づき, 代替経路を提示するアプローチは, これまで十分に確立されていない.

本研究の第二の目的は, 満足度の低下区間を起点として代替経路を提示することである. これは単なる場所の推薦ではなく, 利用者が過去の体験を振り返り, 未来の旅行計画をより良くするための能動的な探索を支援するものである.

本稿では, 上記二点を実現するため, 大規模言語モデル (LLM) を用いて, 旅行記のテキストから①場所・行動・感情を紐付けた旅行体験の可視化, および②体験の質を向上させる代替経路の生成, という2つの手法を統合して提案する.

2. 材 料

本研究では, 「地球の歩き方旅行記データセット」⁽⁵⁾ に収載されている4,500件の国内旅行記のうち, 旅程が記された「旅スケジュール (旅スケ)」付きの3,153件を利用した. 旅スケ付きの旅行記は, 旅行者の訪問場所や行動, 所感が時系列に沿って記述された非構造化テキストデータである. 旅行記の典型的な記述内容には, 旅行者の1日の (旅行) 行動, 訪問場所や風景の描写, 各訪問場所や全旅程についての感想などが含まれる (表1). 旅スケの主な構成要素は, 旅行者が訪問した場所と時間帯である. なお, 写真が付随している旅行記もあるが, 本研究ではテキストのみを分析対象とした.

表1 旅スケジュール付き旅行記の例

【旅行記の例】

宇治川に到着！
紅葉は、ぼちぼちってとこですね～
遊覧船で川下りしながらの昼食タイム。

・・・

この辺では「しじみめし」が名物みたいです。
も～ちろん買って食べました。
速弁の「かに献上」で～す。
美味しそうですね。

【旅スケジュールの例】

1日目 2007年11月17日(土)

xx:xx - xx:xx 自宅

07:30 - 07:50 移動

07:50 - 08:20 徳光 PA (北陸自動車道)
08:20 - 11:40 バス
...
xx:xx - xx:xx 夜ご飯用に、南条 SA で速弁「かに献上」を…
18:20 - 19:30 移動 (バス)
19:30 - 19:35 徳光 PA (北陸自動車道)

3. 手 法

提案手法は、旅行体験の可視化 (3.1 節) と代替経路の生成 (3.2 節) から構成される。

3.1. 旅行体験の可視化

旅行体験の構造を「場所」「行動」「感情」の 3 軸で可視化するため、旅行記からこれら三要素を抽出する。本研究では、大規模言語モデル (GPT-4o) を用いて、これら三要素を抽出する。

3.1.1. 訪問場所および経路の特定 訪問場所および経路の特定は、以下の 3 つのステップからなる。なお、本研究では、「訪問場所」を旅行者が立ち寄った個々の地点として扱い、それらを時系列順に結んだものを「経路」と定義する。

- Step 1: 地名の抽出と曖昧性解消
- Step 2: 地名の座標変換 (ジオコーディング)
- Step 3: 経路データへの変換 (座標系列化)

Step 1: 地名の抽出と曖昧性解消では、旅行記から地名を抽出する。ただし、地名だけでは曖昧性を生じる場合がある (例: 本町, 大久保, 中央公園など)。そのため、まず旅行記全体から訪問した都道府県を判定する。複数の都道府県にまたがる旅行記の場合は、代表的なものを 1 つ判定する。次に、旅行記から訪問場所の地名を抽出し、都道府県名と地名をセットにすることで、曖昧性を軽減する。例えば、「本町から難波まで歩いた。」というテキストからは (本町, 大阪府), (難波, 大阪府) のような地名と都道府県名のセットが得られる。

本研究では、GPT-3.5-turbo に表 2 の、GPT-4o に表 3 のプロンプトを適用することによって、この処理を実現した。

表 2 都道府県名判定のプロンプト

以下の旅行記データから筆者が訪れたと考えられる都道府県を 1 つだけ答えてください。ただし、特定の語句に拘らずに旅行記全体から総合的に判断してください。

Step 2: 地名の座標変換 (ジオコーディング) では、Step 1 で得られた「地名」と「都道府県名」のペアを入力として、それぞれの緯度・経度座標を取得する。本研究では、OpenStreetMap のデータベースを利用するジオコーディングライブラリを用いた。都道府県名をヒントとして付与することで、地名の曖昧性を軽減し、座標特定の精度を向上させる。ジオコーディングが失敗した場合に備え、GPT-4o による地名のみをヒントとした座標推定も行う。このとき、座標だけでなくその座標の推定理由も出力する。ジオコーディングにより、例えば、(本町, 大阪府) → 緯度: 34.683968, 経度: 135.50396 のような変換が行われる。

Step 3: 経路データへの変換 (座標系列化) では、Step 2 で得られた訪問場所の座標データを、入力テキスト中の出現順 (時系列順) に対応付ける。これにより、旅行者の訪問履歴を表す経路データ (座標の系列) が生成される。得られた経路データは、地図上で旅行者の移動経路を可視化するために利用される。

表 3 訪問場所の地名抽出のプロンプト

以下の旅行記のテキストを時系列に沿って分析し、「滞在 (stop)」と「移動 (move)」のイベントを交互に抽出してください。

****抽出ルール:****

- イベントは必ずリスト形式で、`"type"` キーを持つオブジェクトとしてください。
- `"type": "stop"`: ある場所での行動や体験。
 - `"place"`: 地名
 - `"latitude"`, `"longitude"`: GPT による推定座標
 - `"experience"`: その場所での具体的な体験
 - `"reasoning"`: 座標推定の理由
- `"type": "move"`: 場所から場所への移動。
 - `"means"`: 移動手段 (以下のリストから選択)
 - `"experience"`: 移動中の具体的な体験

****移動手段リスト:**** {MOVE_TAGS}

****出力形式の厳守:****

- 必ず JSON 形式のリストとしてください。
- 最初と最後のイベントは、多くの場合 `"stop"` になります。
- `"stop"` と `"move"` は交互に現れるのが基本ですが、テキストに記述がなければ片方が連続しても構いません。

3.1.2. 行動の特定 各訪問場所での具体的な行動を特定するために、既存研究⁽⁶⁾で定義された「徒歩移動」「食事 (飲酒あり)」「名所観光」など約 30 種類の行動タグを用いて、旅行者の行動内容を分類するタスクとして扱う。まず、抽出された訪問場所周辺の旅行記を分析し、旅行者の行動内容を推定する。次に、推定結果のうち、最も関連性の高い行動タグを抽出する。得られた行動タグは、地図上で旅行者の行動を可視化するために利用される。GPT-4o に与えたプロンプトを表 4 に示す。

表 4 行動タグ抽出および感情評価のためのプロンプト

以下のテキストは、旅行中のある「滞在」場所での経験を記述したものです。

このテキストを分析し、以下のステップを同時に実行してください。

- **タグ抽出**:** 提示された「行動」タグリストの中から、テキスト内容に最も関連性の高いタグをすべて選択してください。
 - ****最重要ルール**:** 必ずこのリストに含まれるタグのみを使用し、リストにない新しいタグは絶対に創作しないでください。
- **タグ別感情分析**:** ステップ 1 で選択した各タグについて、そのタグに関連するテキスト部分の感情を個別に

分析し、**-1.0（非常にネガティブ）から 1.0（非常にポジティブ）**のスコアを算出してください。**-ニュートラルな感情は 0.0**とします。

関連性の高いタグが一つもなければ、空のオブジェクト`{}`を返してください。
出力は必ず、キーが「タグ名」、値が「感情スコア」の JSON オブジェクト形式で返してください。

例:
{
 "食事(飲酒なし・不明)": 0.85,
 "景色鑑賞": 1.0
}

3.1.3. 感情価の評価 各訪問場所における体験の質を評価するため、旅行記の感情評価を行う。具体的には、訪問場所に関連する体験が記述された文章を対象とし、その感情価を-1（最もネガティブ）から 1（最もポジティブ）の範囲でスコアリングする。感情価の評価には、ワンショットプロンプティング手法を採用した。GPT-4o に与えたプロンプトを表 4 に示す。

3.1.4. 可視化 各旅行記から抽出した経路データの「場所（座標）」「行動（アイコン）」「感情価（スコア）」の三要素を統合し、インタラクティブな地図上に可視化する。訪問場所は時系列に沿って線で結ばれ、旅行者の移動経路を表す。感情価はヒートマップとして地図上に重畳表示され、ポジティブおよびネガティブな体験が発生した地域を直感的に把握できるようにする。さらに、各行動アイコンをクリックすると、該当する旅行記の原文、感情価および GPT による座標の推定理由を確認できるポップアップを表示する。

3.2. 代替経路の生成

次に、可視化によって特定される「満足度のボトルネック（ネガティブな体験）」を起点とし、未来の計画立案を支援する代替経路を生成する。この生成は、「もしあの時、別の行動をとっていたなら」という代替の可能性を LLM によって生成し、探求するものである。具体的には、まず、前節の感情評価の結果から、感情価が一定の閾値を下回るネガティブな体験区間を特定する。次に、該当する旅行記のテキストから、そのネガティブな体験が記述された箇所

（中間部分）をマスク（削除）する。最後に、マスクされたテキストの前後の文脈（すなわち、ネガティブな体験の「前」と「後」の行動）を LLM（GPT-4o）へのプロンプトとして入力する。これにより、LLM は前後の文脈を自然に繋ぎつつ、元のネガティブな体験とは異なる「あり得たかもしれない」代替の旅行体験ストーリーを生成する。生成された代替ストーリーに対しても同様の可視化プロセスを適用し、元の経路と比較可能な形で地図上に提示する。

4. 結 果

本節では、提案手法による旅行体験の可視化結果と、ネガティブな体験に基づく代替経路の生成結果について述べる。

4.1. 旅行体験の可視化結果

提案手法を用いて旅行記のテキストを分析した結果、旅行行程における「場所」「行動」「感情」の三要素を統合して可視化した。データセットから抽出された 4,314 件の行動のうち、ネガティブなもの（ここでは、感情価が 0 未満のものとする）は 62 件（1.44%）であった。地名から特定困難なものを除くと、ネガティブな体験が見られる場所は、地域別では九州・沖縄エリアが最多の 8 件、次点で関東エリアが 6 件となった。特性別では、最も多いのは「自然（山・海・皮・公園等）」(7 件) である。次に多いのは「宿泊施設」、ついで「交通関連施設」「ランドマーク・史跡」であった。

図 1 に、ある旅行記の可視化結果の一例を示す。地図上には訪問経路が時系列に沿って描画され、各地点での行動がアイコンで示されている。感情価はヒートマップとして重畳表示されており、この例ではほとんどの訪問場所で感情価が高い一方で、仁和寺での体験(図中左上)の感情価が低いことが確認できる。このように、提案手法によって、旅行者が行程の「どこで」「何をして」「どのような感情を抱いたか」を直感的に把握できることが示された。

4.2. 代替経路の生成結果

次に、満足度のボトルネック（ネガティブな体験）を起点とした代替経路の生成結果を示す。表 5 に、分析対象とした旅行記の概要を示す。この旅行記では、特に「仁和寺」訪問時に「御室桜が見頃を過ぎていた」というネガティブな体験が記述されていた。

表 5 ネガティブな体験を含む旅行記の構造化の例。概要は旅行記のテキストを LLM が要約したもの。

場所	行動	感情価	概要
東寺	軽食(カフェなど) 景色鑑賞	感情価: 0.95	満開の桜の中、抹茶ソフトを食べながら春の京都を満喫。
嵐山（渡月橋）	名所観光 散歩	感情価: 1.00	風景を楽しみ、天龍寺を拝観。
仁和寺	景色鑑賞	感情価: -0.50	御室桜を目的に訪問したが、見頃を過ぎており葉桜の状態。
龍安寺（石庭）	景色鑑賞	感情価: 1.00	過去（中学生時）の訪問時とは異なり、石庭の魅力を理解し、癒しの効果を肯定的に評価。
金閣寺	景色鑑賞 軽食(カフェなど)	感情価: 0.90	池に映る金閣寺の景観を高く評価。
建仁寺	景色鑑賞	感情価: 0.50	寺社巡りの最後の訪問地として、庭園を鑑賞しながら旅行全体の内省を行った。

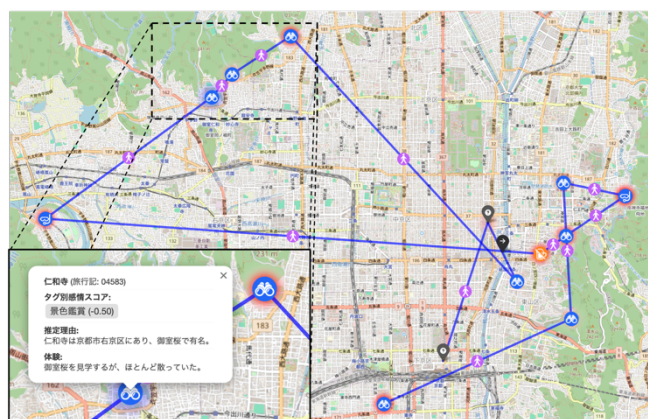


Figure 1. An example visualization of a Kyoto City travelogue containing negative experiences. A negative experience at Ninna-ji Temple can be seen in the pop-up. The blue line indicates the route, the icons represent activities at visited locations, and the color surrounding the icons indicates positive (red) and negative (blue) experiences.

表 6 ネガティブな体験を含む旅行記の本文（抜粋）と生成された代替テキスト

<p>【直前】最初は、嵐山の渡月橋です。この日はスケジュールを少しきつめに設定。あまりのんびりせずに、テキパキ行動します！天龍寺を拝観し、竹林の道へと抜けました。竹が青々としていて、ホントにキレイです！ちょっと神秘的な感じもしました。</p>
<p>【直後】修学旅行以来の龍安寺です。中学生だった当時は、石庭の魅力が分からず、石の何がいいんだろうと思っていましたが、大人になってから改めて訪れると、心から癒される感じがしました。・・・和みます。</p>
<p>【中間】昼前に仁和寺に到着。こちらは桜が有名なお寺ということで、3日目のメインスポットだったのですが、残念ながら名物の御室桜は半分以上散っていて、葉桜になっていました…。3日目ではなく1日目に来ていたら、もう少し咲いていたかな。</p>
<p>【出力】ChatGPT (GPT5)</p> <p>竹林を抜けたあとは、そのまま野宮神社へ立ち寄りしました。境内に入ると、苔むした黒木の鳥居がひっそりと佇んでいて、周囲は竹林のざわめきしか聞こえません。小さな社ですが、縁結びや学業成就のご利益があると聞き、苔の上に手を合わせてみました。ちょうど和装の前撮りをしているカップルがいて、朱色の衣装と竹の緑のコントラストがとても絵になっていました。そのあと、神社近くの茶店で「よもぎ団子 (3本入りで400円)」をいただきました。焼き立ての団子からは香ばしい香りが立ちのぼり、口に入るとほのかな苦みと甘いみたらしのタレが絶妙で、思わずもう一本追加しそうなったほど。値段も観光地にしては良心的で、大満足でした。</p>

生成された代替ストーリーは、「仁和寺」訪問（ネガティブな体験）の代わりに、「野宮神社」訪問と「茶店での間食」

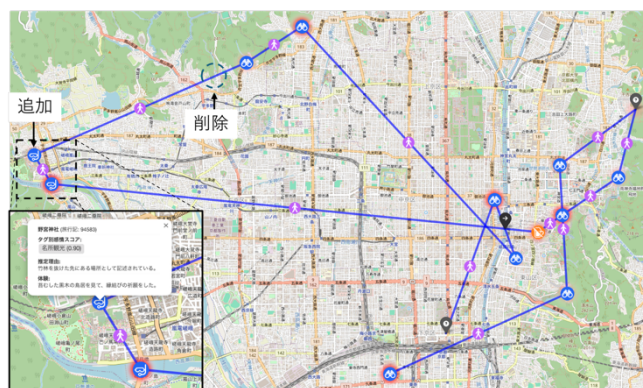


Figure 2. A visualization result of the generated alternative route. The negative experience at Ninna-ji Temple has been removed, and a positive experience at Nonomiya Shrine has been added.

というポジティブな体験を提案している。この代替ストーリーは、直前の「竹林の道」から直後の「龍安寺」への地理的・時間的な文脈を保ちつつ、元のネガティブな体験とは異なる「あり得たかもしれない」旅行の可能性を示している。この結果は、元の経路では仁和寺で感情価が低下したが、代替経路では野宮神社周辺で高い感情価が推定されることを示唆している。

5. 議論

5.1. ネガティブな体験の分析

本研究で扱ったネガティブな体験は、当日の天気や時期のずれによって期待していた景色や観光ができないことに起因するものが大半であった。名所観光における期待は寺のような歴史的建造物に対するものが多く、時間の経過による景観の変化や保護のための設備が景観を損ねていた例、工事や焼失のために観光ができない状態であったケースが多く見られた。

5.2. LLMによる観光体験の提案の評価について

本研究ではマスクしたネガティブな部分を LLM で生成し、元の体験に比べてポジティブな内容を提案することを目指したが、LLM によるシミュレーションが実際にポジティブな体験を提案することができるか検証するため、以下のような実験が考えられる。元の旅行記においてポジティブな体験をマスクし、正解データとおく。マスクした部分を LLM で生成し、正解データに近い体験を提案できるかどうかを検証する。実際のポジティブな旅行体験を LLM で提案できることが示されれば、ネガティブな部分に対してもポジティブな体験を提案できることが示されるといえる。

5.3. 今後の展望

本研究で行った旅行体験の可視化および代替経路の提案について、今後以下のような課題があり、研究を推進する予定である。

可視化：本研究で用いたデータセットには「旅行記」ファイルの他に「旅スケジュール」ファイルも含まれているが、旅スケの内容と旅行記の内容の統合が困難だったため、本研究では可視化結果には旅行記の内容のみ反映した。そのため、旅行記内で言及されることの少ない移動手段が不正確な部分が多く見られた。今後は旅スケジュールの内容

も活用することで、より高い精度の体験抽出および可視化を目指す。

評価：作成した可視化結果がユーザーにとって旅行体験の全体像や満足度のボトルネックの理解に役立つかどうかを確かめるため、評価を行う。具体的には、被験者に対し、従来のテキストのみの旅行記と、本手法による可視化結果を提示し、特定の旅行記に対する読解タスクを遂行してもらうことを予定している。この結果、タスクの完了時間や正答率を計測ができ定量的な評価が可能となる。同時に、アンケートおよび半構造化インタビューなども実施予定である。

- 旅行行程全体の流れや、各地点での体験の質（ポジティブ/ネガティブ）を理解するのに役立ったか
- 感情価を示すヒートマップや、行動を示すアイコンが、旅行者の体験内容を直感的に伝えていたか
- 満足度が低下した箇所や、その原因を特定しやすかったか
- 地図上をクリックして詳細なテキストや評価理由を確認する機能が、体験の深堀りに役立ったか
- テキストのみの場合と比較して、体験の分析が用意であったか

同様に代替経路に関しても、定量的な評価を以下のように実施予定である。

- 文脈的一致性：生成された代替ストーリーが、ネガティブな体験の直前と直後の文脈を破綻なく自然に繋いでおり、「あり得る可能性」として立とうかどうかを LLM-as-a-Judge または人手により評価する。
- 教師データからの逸脱度：モデルが元のネガティブな体験（教師データ）とは異なる、新しい可能性を探求できているかを ROUGE スコアにより評価する（この評価では、スコアが低いほど元のテキストとは異なる、新規のシナリオを生成できていると解釈する）。
- 生成の多様性：モデルが「あり得る可能性」を 1 パターンだけでなく、多様に提示できているかを Self-BLEU により評価する（スコアが低いほど、生成結果どうしが似ておらず、様々なパターンの出力をしていると解釈する）。

6. おわりに

本研究は、Web 上に蓄積された個人の旅行記データを分析し、旅行体験の理解（可視化）と計画立案（代替経路の提案）を支援するための新たなフレームワークを提案した。本研究で得られた主要な知見は以下の通りである。

第一に、大規模言語モデル（LLM）を活用することで、旅行記に含まれる「場所（経路）」「行動（体験）」「感情（体験の質）」という三要素を抽出し、これらを時系列に沿って地図上に統合・可視化できることを示した。これにより、感情から満足度の「ボトルネック（ネガティブな体験）」を特定でき、さらにボトルネックの場所の情報も得ることができる。このボトルネックのような文脈に依存したイベントの特定は従来の困難であったが、本研究で初めて実現できたと考えている。

第二に、可視化によって特定された「ボトルネック」を起点として、「もしあの時、別の行動をしていたら」という文脈に基づいた代替の旅行体験ストーリーを LLM に生成さ

せる、新たなプランニング支援の手法を提案した。これは、従来の地点推薦システムとは異なり、利用者が過去の体験を振り返り、未来の計画をより良くするための能動的な探索を支援するアプローチである。

本研究の成果は、膨大な旅行体験の記録を単なる過去のアーカイブとしてではなく、未来のより良い旅行をデザインするための実践的な知見へと転換する可能性を示した点にある。

今後の課題としては、提案した可視化手法が、利用者の旅行体験の理解やボトルネックの把握に実際にどの程度寄与するかを、ユーザビリティテストを通じて定量的に評価する必要がある。また、LLM によって生成される代替経路に関しても、その文脈的な整合性、生成される体験の多様性、そしてそれが実際に元のネガティブな体験を上回るポジティブな体験を提案できているかを、定量的・定性的に評価する手法の確立が求められる。これらの検証を通じ、本フレームワークの実用性をさらに高めていきたい。

謝 辞

本研究の一部は、JSPS 科研費 22H03648, JST SICORP JPMJSC23C6 の支援を受けたものである。

文 献

- (1) Chu M, Chen Y, Yang L and Wang J: Language interpretation in travel guidance platform: Text mining and sentiment analysis of TripAdvisor reviews. *Front. Psychol.* 13:1029945. 2022. doi: 10.3389/fpsyg.2022.1029945
- (2) Ma, Y., Wang, Y., Xu, G., and Tai, X.: Multilevel Visualization of Travelogue Trajectory Data. *ISPRS International Journal of Geo-Information*, 7(1), 12. 2018. <https://doi.org/10.3390/ijgi7010012>
- (3) Matsumoto Y., "Tourism Information Analysis of SNS Data Using Text Mining," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 29 No. 4, pp. 829-837, 2025
- (4) 山岸立, 馬強: 個人適応型観光のためのユーザ体験推薦, 日本データベース学会和文論文誌, Vol. 20-J, Article No.15, 2022 年 3 月
- (5) 大内啓樹, 進藤裕之, 若宮翔子, 松田裕貴, 井之上直也, 東山翔平, 中村哲, 渡辺太郎: 地球の歩き方旅行記データセット, 言語処理学会 第 29 回年次大会 発表論文集, pp. 2920-2924, 2023 年 3 月
- (6) 中岡明義, 若宮翔子, 荒牧英治: 行動分類のためのコーパス構築と行動分析への応用, 言語処理学会 第 31 回年次大会 発表論文集, pp. 2475-2480, 2025 年 3 月