



地球の歩き方 旅行記データセット

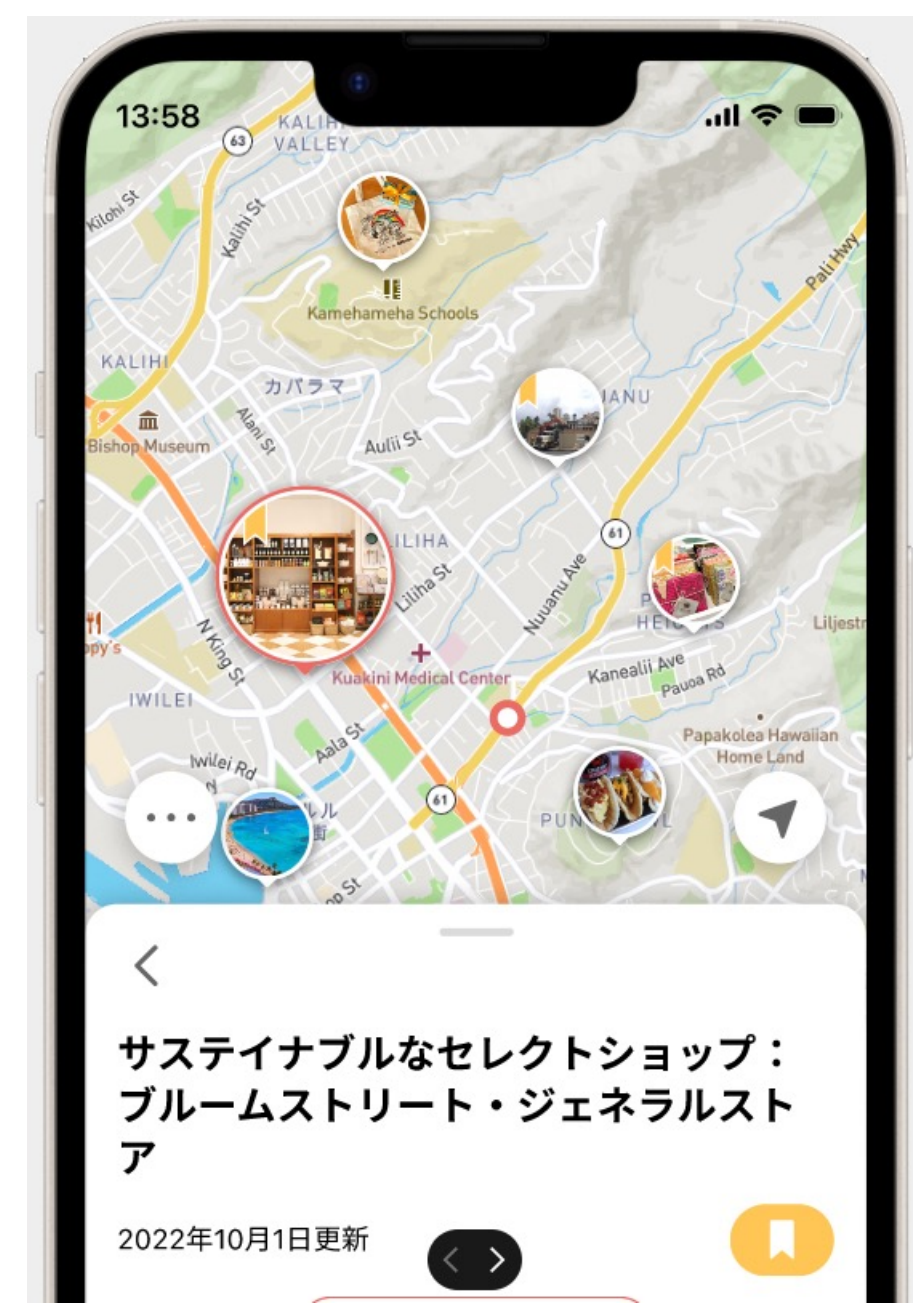
(株)「地球の歩き方」について

- ◆ 1979年に創刊した旅行ガイドブック「地球の歩き方シリーズ」を中心とする出版事業と、インバウンド・アウトバウンドの旅行マーケティング事業を行っております。2021年より株式会社GAKKENの子会社として再始動しました。
- ◆ 「旅」の新しい文化的価値創出を広くみなさまに提案することを目的にし、従来のガイドブックに加え、企業間・自治体コラボ、デジタルメディアほか複合的な事業展開にも注力しています。



(株)「地球の歩き方」での位置データとUGCコンテンツの活用

- ◆ 地球の歩き方ウェブでは、長年のガイドブック制作で培ってきたノウハウをデジタル領域に再現することをテーマに海外の豊富な登録物件のコンテンツ化を進めています。将来的な位置データを用いたサービスなどにつなげて行く予定です。



- ◆ また、サイト内では、会員限定サービスとして、世界の旅行者同士の情報交換、会員間チャットを試験的に提供しています。ガイドブックへのフィードバックや、地図サービスと掛け合わせたサービスを見込んでいます。



データセットの内容

旅行記

会津若松へ向かう磐越西線の接続を考慮して選んだやまびこ 203号はE5系での運転でした。何度も乗っているE2系よりも座席が広く感じ、快適な移動でした。

会津若松駅から快速あいづ4号に乗り、郡山へ向かいました。会津若松は晴れ間がありましたが、山を上につれて雲が増えて行き、途中から雨が降り出しました。天気予報通りでしたが、今回の旅行は暖かい2日間で移動時間を除いて雨に降られることがなかったのはラッキーでした。

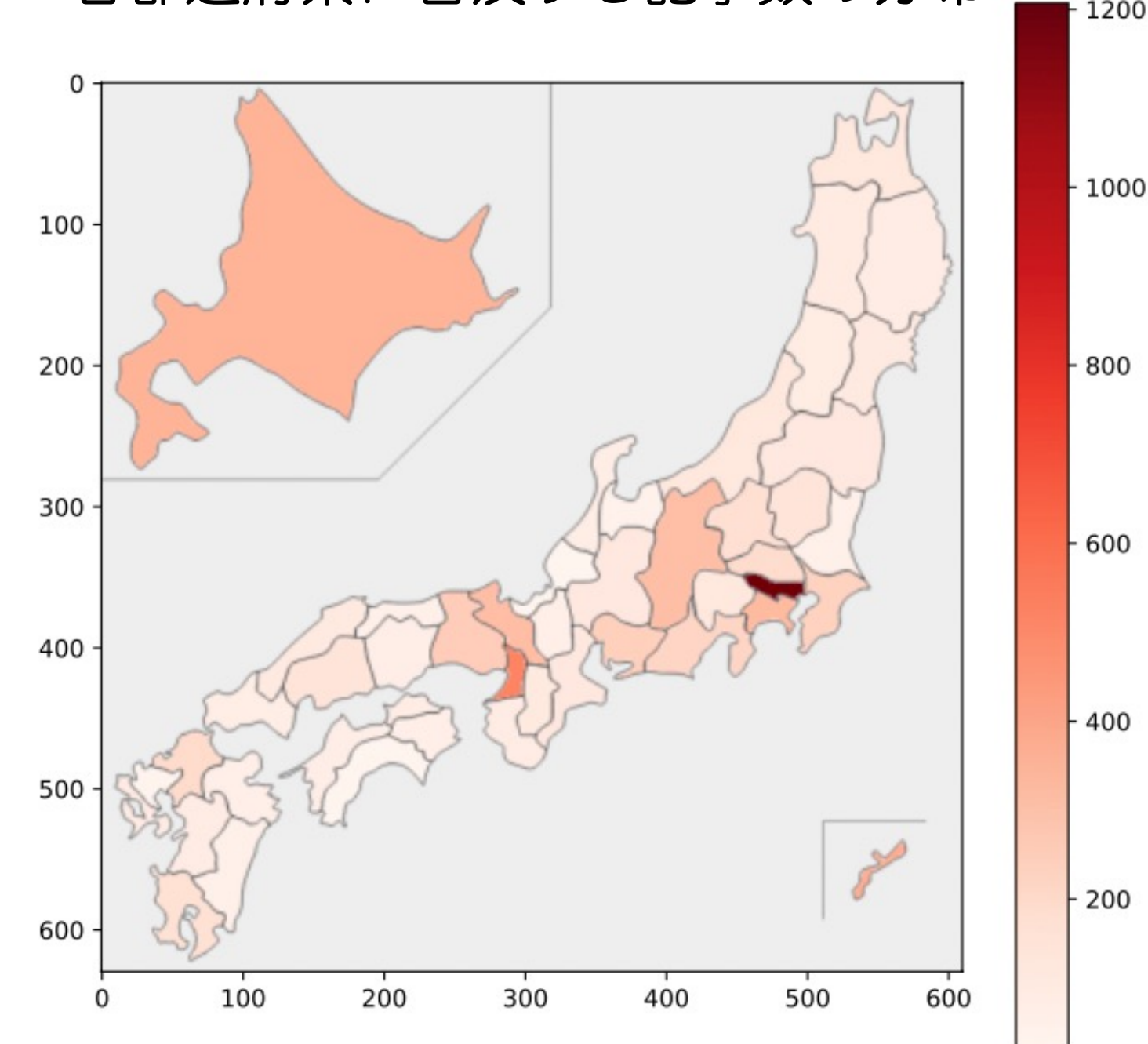
旅スケジュール

1日目	2021年10月15日(金)
05:40 - 05:50	自宅
05:50 - 05:53	最寄駅
...	
15:33 - 23:59	庄助の宿 瀧の湯
2日目	2021年10月16日(土)
00:00 - 09:25	庄助の宿 瀧の湯
09:25 - 09:32	東山温泉入口 (瀧の湯前) バス停
...	
17:33 - 17:38	最寄駅
17:38 - 17:45	自宅

データセットの特徴

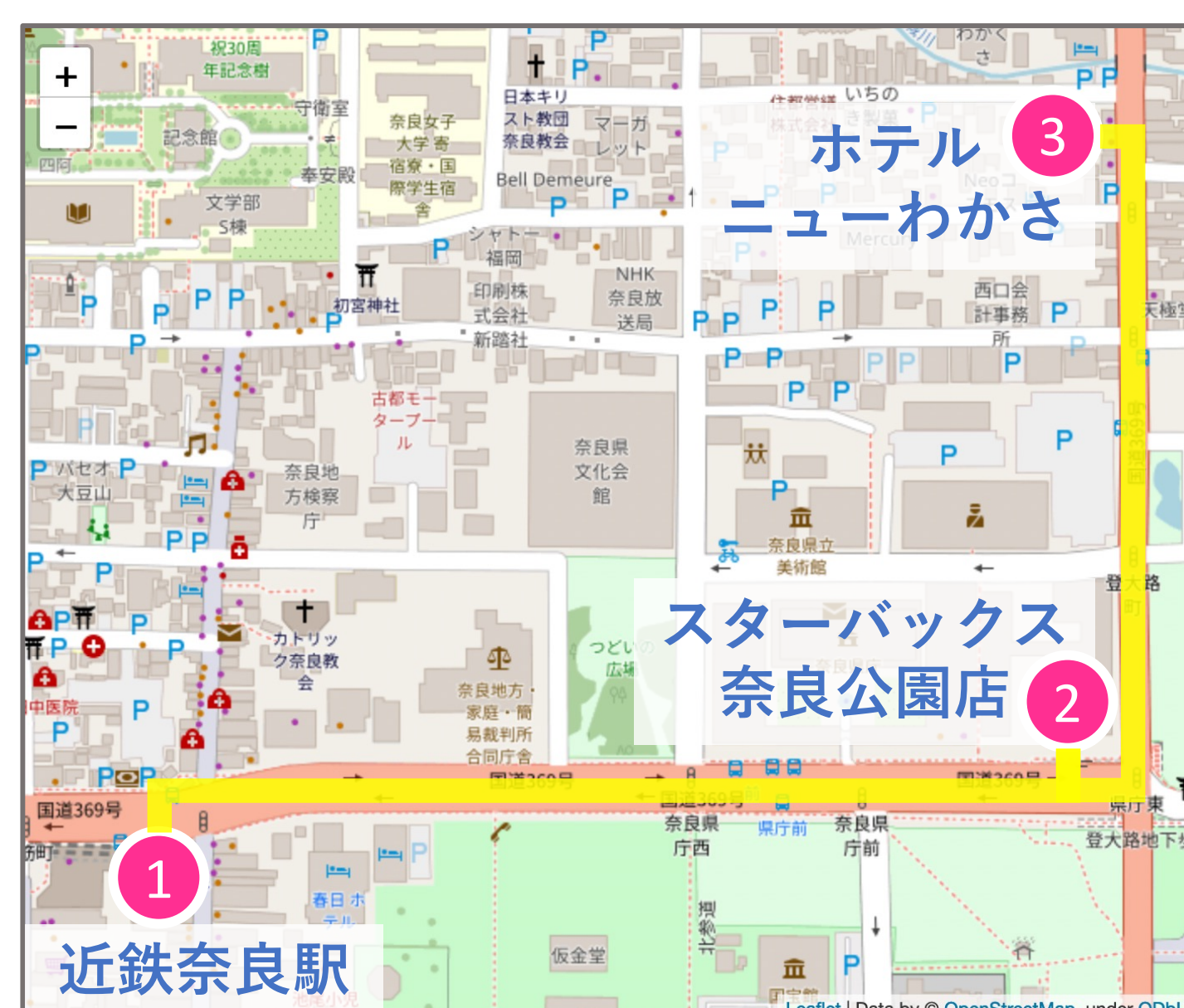
- ◆ 国内旅行 4,500 記事
- ◆ 海外旅行 9,500 記事
- ◆ 計 3,000万文字 超えの巨大データセット
- ◆ 全都道府県をカバー
- ◆ 200以上の国と地域をカバー

各都道府県に言及する記事数の分布



期待される活用法

- ◆ 旅行者の移動行動分析
- ◆ 観光地のトレンド分析
- ◆ 穴場観光スポット発掘
- ◆ 旅行計画/推薦アプリ
- ◆ 地域振興への応用



実際の活用事例

- ◆ 奈良先端大(NAIST) 自然言語処理研究室の研究成果
- ◆ 本データセットをもとにAIを構築
- ◆ 入力文から自動的に地名・施設名を抜き出し、地図上の位置と紐付けて可視化

