

文章から地名を自動的に見つけ出す

GeoNLP: 自然言語文のジオタギングのためのソフトウェア環境

どんな研究？

- 文章中の地名を自動的に抽出。
- 複数候補の中から地名を決定。
- 地図上に表示して把握。
- いろいろな文章の地名を統合。

何ができる？

- いまどこで何が起こっている？
- ある場所のニュースを追跡。
- 他のデータと組み合わせ分析。
- 歴史的な地名で古文書进行分析。

GeoNLP ソフトウェア（地名抽出・解決ツール） 2013 年夏に公開予定！

地名の抽出と解決



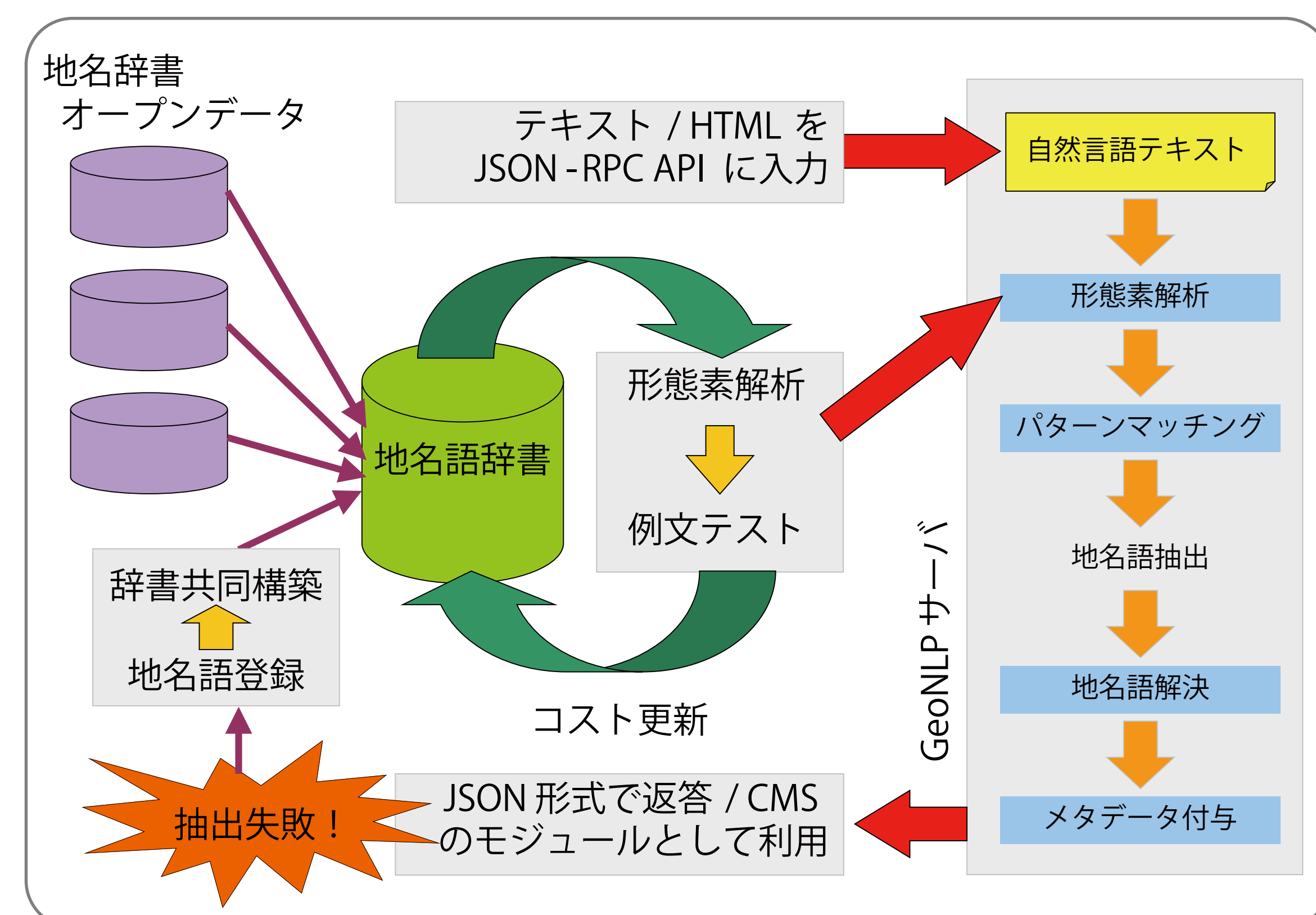
横浜は雪だよ。川崎はどう？



横浜雪なんだ！川崎は雨だよ。

神奈川県横浜市と神奈川県川崎市の話？
青森県横浜町と福岡県川崎町の話？

複数の候補があると一意に決まらない。
他の手がかりを使って決める必要がある。
例えば、地名が同じグループか、同じ種類かなどをチェック。



1. ウェブサービスの提供
2. コマンドラインツールの配布

GeoNLP データ（地名語辞書共有サイト）

2013 年夏に公開予定！

共有のメリット

1. 地名語辞書として必要な項目名を共通化。
2. 地名語 CSV を辞書サーバにアップロードして共有。
3. 辞書サーバがユニーク ID を自動付与して管理。
4. 地名検索・地図化インタフェースを利用可能。
5. GeoNLP ソフトウェアと辞書を同期して地名解析。
6. 政府オープンデータも地名語辞書に変換して活用。
7. Geo Linked Open Data として他のデータと連携。

みんなで協力して、地名辞書を充実させよう！

（自然地名）岬や島の地名、（施設名）避難所の名前、（POI）レストランの名前、（地域情報）秋葉原のスポット名、など。

本研究は、相良毅（株式会社情報試作室）およびトライアックス株式会社の協力を得ています。
科学技術振興機構（さきがけ）および地球環境情報統融合プログラム (DIAS) の支援を受けました。

GeoNLP を取り巻くエコシステム

