

歴史的行政区域データセットβ版 をはじめとする幾何データ共有 サイト「Geoshape」の構築

北本 朝展^{1,2} 村田 健史³

1. ROIS-DS人文学オープンデータ共同利用センター
2. 国立情報学研究所
3. 情報通信研究機構

<http://geoshape.ex.nii.ac.jp/>

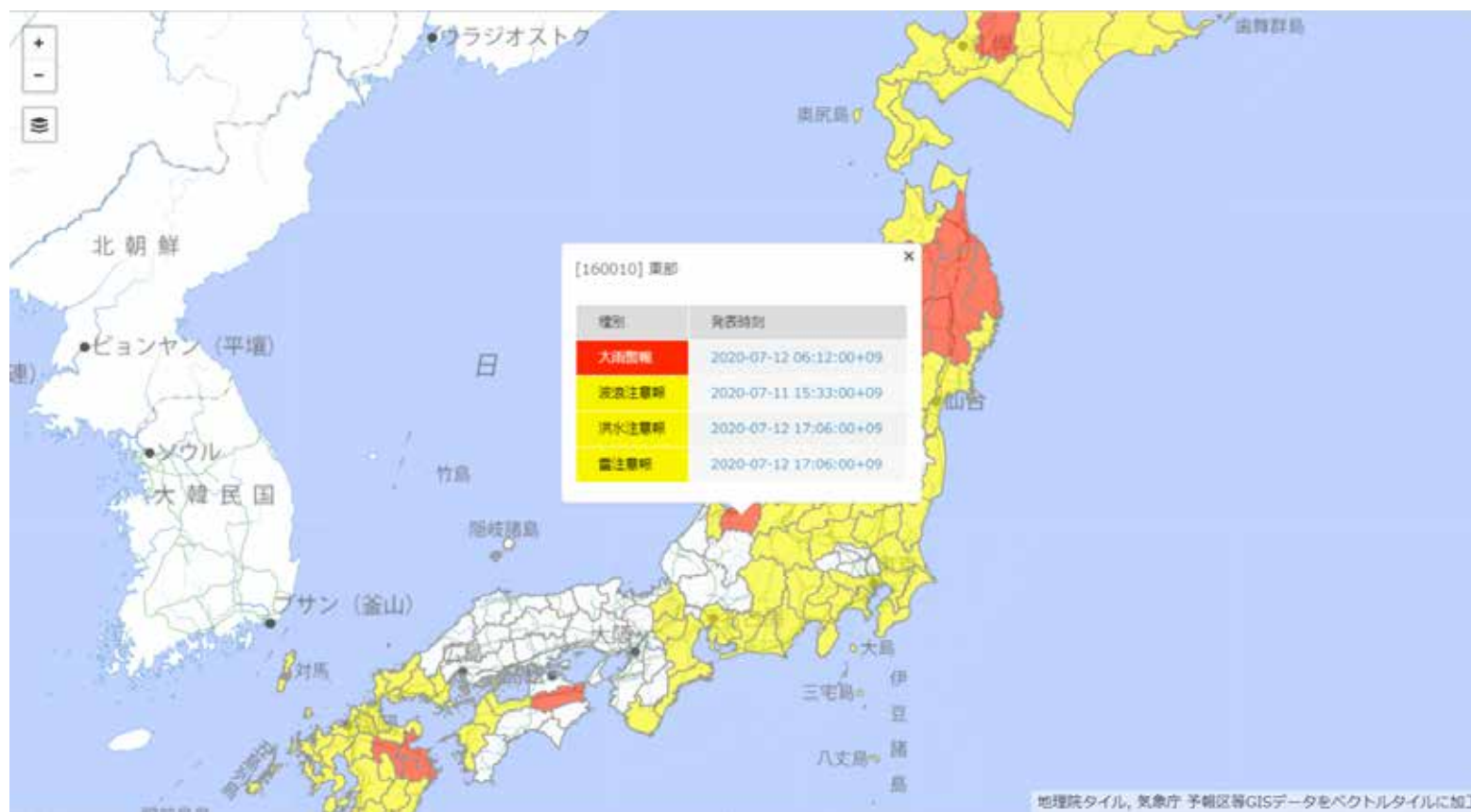
Geoshape公開データ

<http://geoshape.ex.nii.ac.jp/>

1. **歴史的行政区域データセットβ版**：86,107件
 2. **国勢調査町丁・字等別境界データセット**：
219,271件
 3. **気象庁防災情報発表区域データセット**：
10,068件
 4. **国土数値情報河川データセット**：28,192河川
- データセットの可視化、ベクトルタイル地図によるコロプレス地図、各種検索機能などを提供。
 - **気象庁防災情報XML**等のサービス展開の基盤。

気象警報最新マップ

<http://agora.ex.nii.ac.jp/cps/weather/warning-map/>



気象庁防災情報発表区域データセット：Zoom 5は府県予報区等、Zoom 6-7は一次細分区域等、Zoom 8は市町村等をまとめた地域等、Zoom 9-13は市町村等（気象警報等）

国土数値情報河川データセット

<http://geoshape.ex.nii.ac.jp/river/>



利根川水系 (830303) | 国土数値情報河川データセット :
<http://geoshape.ex.nii.ac.jp/river/resource/830303/>

地図データと幾何データ

1. **地図データ**：緯度経度など数学的に定義された座標系の上で、各種の情報を統合する。伝統的なGIS (Geographic Information Systems)等も、緯度経度などの座標系が基盤となる。
 2. **幾何データ**：地理的なエンティティを列挙し、ID単位で各種の情報を統合する。エンティティの特定（同一性の判断）、固有IDの付与、属性情報の整備などが重要な課題。
- 近年のウェブ地図技術の進化を取り入れれば、効率的なデータ共有が実現できる。

データ記述形式の進化

1. **シェープファイル (shp)** : バイナリ形式。複数のファイルを組み合わせる必要あり。
2. **GML (Geography Markup Language)** : XML形式。仕様が巨大で重く、気軽に使えない。
3. **KML (Keyhole Markup Language)** : XML形式。Google Earthで普及。GMLよりは使いやすい。
4. **GeoJSON** : JSON形式。ウェブサービスで一般的なJSON形式が嬉しい。
5. **TopoJSON** : JSON形式。複数のポリゴンで境界を共有でき、座標情報も圧縮できる。この工夫は、特にコロプレス地図に効果を発揮する。

データ配信形式の進化

1. **OGC WMS (Web Map Service)** : 動的生成方式。緯度経度範囲をリクエストし画像を生成。
2. **Google Maps** : タイル方式。地球表面を階層的に分割したタイルごとに画像をオフラインで生成。キャッシュを活用して負荷軽減。その後Tile Map Service (TMS) として広く普及。
3. **ベクトルタイル** : タイル方式。描画に必要な情報を配信し、画像生成はクライアント側。
4. **バイナリベクトルタイル** : Mapbox Vector Tile 形式。Protocol Buffersを用いてバイナリ圧縮。

Geoshapeの現状

1. **エンティティ地図**は、初期はGeoJSONが中心。その後はTopoJSONを追加し、GeoJSONとTopoJSONの両方を提供。
2. **コロプレス地図**は、初期はTopoJSONを使い、その後はバイナリベクトルタイルに移行。
3. **元データの大多数はシェープファイル**。その**幾何形状**を、各種の**オープンソースツール**で、上記の形式に変換して公開。
4. シェープファイルの**属性情報**を取り出し、検索や表示に活用するシステムを提供。

歴史的行政区域データセットβ版

<http://geoshape.ex.nii.ac.jp/city/>



1920年から
2019年まで、
27時点の行政
区域の表示を
オン・オフで
きる。

現在の行政区
域と重なる過
去の市区町村
などの情報も、
計算の上掲載。

市町村大合併と地域感覚

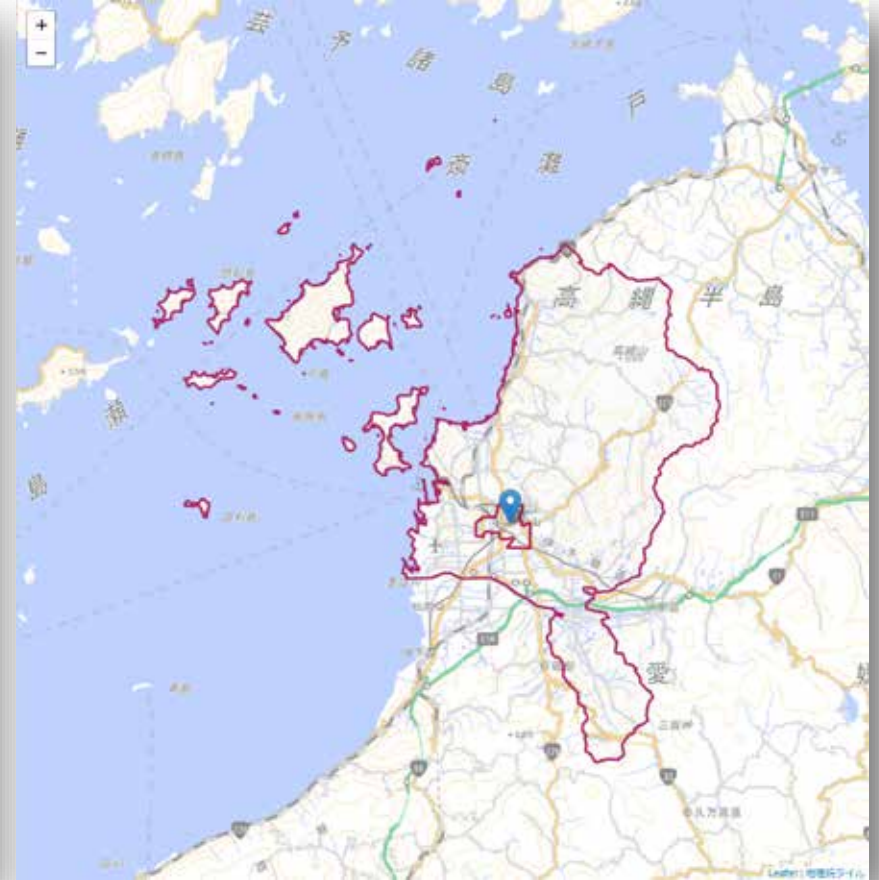
1. **明治の大合併**（約300～500戸規模）：71,314
（M21）→ 15,859（M22）
 2. **昭和の大合併**（約8,000人規模）：9,868
（S28）→ 3,472（S36）
 3. **平成の大合併**：3,218（H14）→ 1,727（H22）
- 幾度もの大合併を経て大きくなり過ぎた市区町村では、**行政区域内部での地域差が拡大**。
 - **旧市区町村名は住民の日常生活や歴史の記憶とリンク。災害情報等も感覚的に理解できる。**

参考：<https://www.soumu.go.jp/gapei/gapei2.html>

合併で拡大した市区町村の例



山形県鶴岡市



愛媛県松山市

データソースの問題

https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-N03-v2_4.html

- 国土交通省が作成する「**国土数値情報**」は、過去の行政区域に関する貴重なデータだが...
- **問題1**：シェープファイル形式のため、ウェブサービスの利用には適していない。
- **問題2**：時間でスライスしているため、市区町村の変遷を時系列でたどりにくい。
- **問題3**：市区町村に固有のIDがないため、どの市区町村が同じなのか、判断しづらい。
- **国土数値情報データをもっと有効に活用したい。**

歴史的行政区域データの要件

1. 過去から現在に至る**市区町村の変遷**を、追跡可能なデータセットが欲しい。
2. 同一市区町村のデータを**固有のIDで統合**し、時系列データとして使いたい。
3. 境界データに加えて**代表点**を選定し、地理情報としての有用性を高めたい。
4. **ウェブ地図での利用**に適したデータ形式に変換し、使えるようにしたい。

データソース

- 行政区域データ

- 国土数値情報ダウンロードサービスで入手できるXML形式のデータ (N03) 1920年～2019年

- 市区町村役場データ

- 国土数値情報ダウンロードサービスで入手できるXML形式のデータ (P34) 2014年

- 市町村役場等及び公的集会施設データ

- 国土数値情報ダウンロードサイトで入手できるXML形式のデータ (P05) 2010年

- 廃置分合等情報

- 政府統計の総合窓口 (e-Stat) の「廃置分合等情報」を探すから入手できるCSV形式のデータ 1970年4月1日以降

- 標準地域コード

- 政府統計の総合窓口 (e-Stat) の「標準地域コードを探す」から入手できるCSV形式のデータ

市区町村IDの付与

13B0010022	東京都野増村/東京府野増村
13B0030001	東京都西多摩郡霞村/東京府西多摩郡霞村
13B0030002	東京都西多摩郡吉野村/東京府西多摩郡吉野村
13B0030005	東京都西多摩郡古里村/東京府西多摩郡古里村
13B0030006	東京都西多摩郡戸倉村/東京府西多摩郡戸倉村
13B0030007	東京都西多摩郡三田村/東京府西多摩郡三田村
13B0030008	東京都西多摩郡小河内村/東京府西多摩郡小河内村
13B0030009	東京都西多摩郡小宮村/東京府西多摩郡小宮村
13B0030010	東京都西多摩郡小曾木村/東京府西多摩郡小曾木村
13B0030013	東京都西多摩郡成木村/東京府西多摩郡成木村
13B0030014	東京都西多摩郡西秋留村/東京府西多摩郡西秋留村
13B0030015	東京都西多摩郡西多摩村/東京府西多摩郡西多摩村
13B0030016	東京都西多摩郡青梅町/東京府西多摩郡青梅町
13B0030019	東京都西多摩郡増戸村/東京府西多摩郡増戸村
13B0030020	東京都西多摩郡多西村
13B0030021	東京都西多摩郡大久野村/東京府西多摩郡大久野村
13B0030022	東京都西多摩郡調布村/東京府西多摩郡調布村
13B0030025	東京都西多摩郡東秋留村/東京府西多摩郡東秋留村
13B0030028	東京都西多摩郡氷川町
13B0030030	東京都西多摩郡平井村/東京府西多摩郡平井村
13B0070001	東京都南多摩郡稲城村/東京府南多摩郡稲城村
13B0070002	東京都南多摩郡横山村/東京府南多摩郡横山村

1. 1968年以降は標準的な市区町村コードが存在。
2. 1968年以前は、名称の連続性などを根拠に固有のIDを付与（網羅性無保証）。
3. 2種のID形式に分割。
 1. 12345**A**6789
 2. 12**B**3456789

代表点の付与



1. **現在**：市区町村役場の場所を代表点に選定。
2. **過去**：ポリゴンに含まれる公的集会施設の場所（支所等）を代表点に選定。
3. 重心は住民にとっての代表点ではないため、可能な限り**公的なPOI（Point of Interest）**を選ぶ。
4. 合併時は、**現在から過去に向けての連続性**を考慮。

包含関係・隣接関係など

1. 現在の市区町村行政区域と重なる過去の市区町村
 2. 過去の市区町村行政区域と重なる現在の市区町村
 3. 現在の町丁・字等境界と重なる過去の市区町村
 4. 隣接行政区域
 5. 近隣行政区域
- ・ポリゴン・ポイントに対する幾何計算から導出。

β版の公開・更新

- **2017年1月公開**。以降、毎年更新。2020年も8月に更新予定。
- **2019年8月現在**、行政区域件数 = 16,459件、行政区域境界データ件数 = 86,107件、統合境界データ件数 = 10,822件を提供。
- 「β版」としている理由は、**データに各種の誤りが残っている**ため。
- **データの修正は本来なら国土交通省が行うべき**だが、反応は鈍い。どうしたものか。

国勢調査町丁・字等別境界データセット

<http://geoshape.ex.nii.ac.jp/ka/>

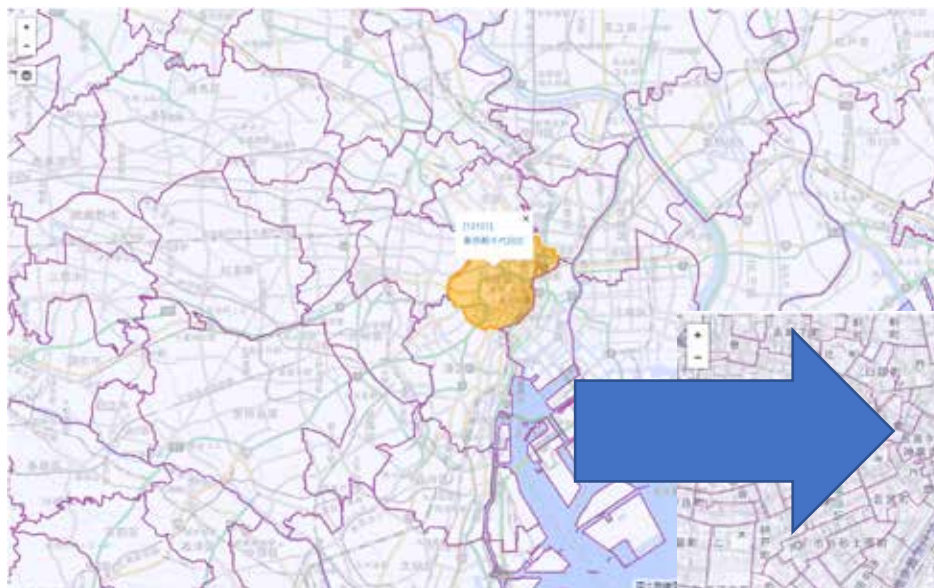
1. 国勢調査実施毎（5年毎）に設定した調査区の境界を中心としたデータセット。
2. 2015年版のデータでは、町丁・字等情報 = 219,271件を提供。
3. 元のデータソースはe-Stat。これをNICTが加工し、NIIがさらに加工。
4. **時間は1点のみだが、空間解像度は市区町村行政区域よりも細かい。**

<http://geoshape.ex.nii.ac.jp/ka/resource/13/13201050001.html>



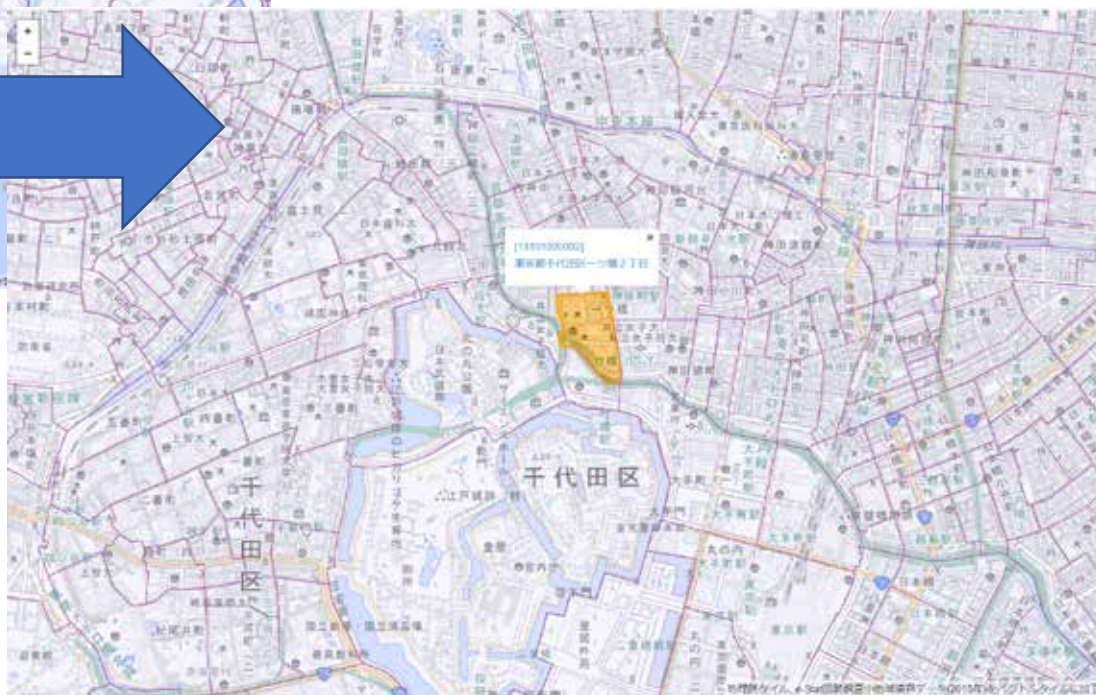
市区町村と町丁・字の統合

<http://geoshape.ex.nii.ac.jp/vector-adm/>



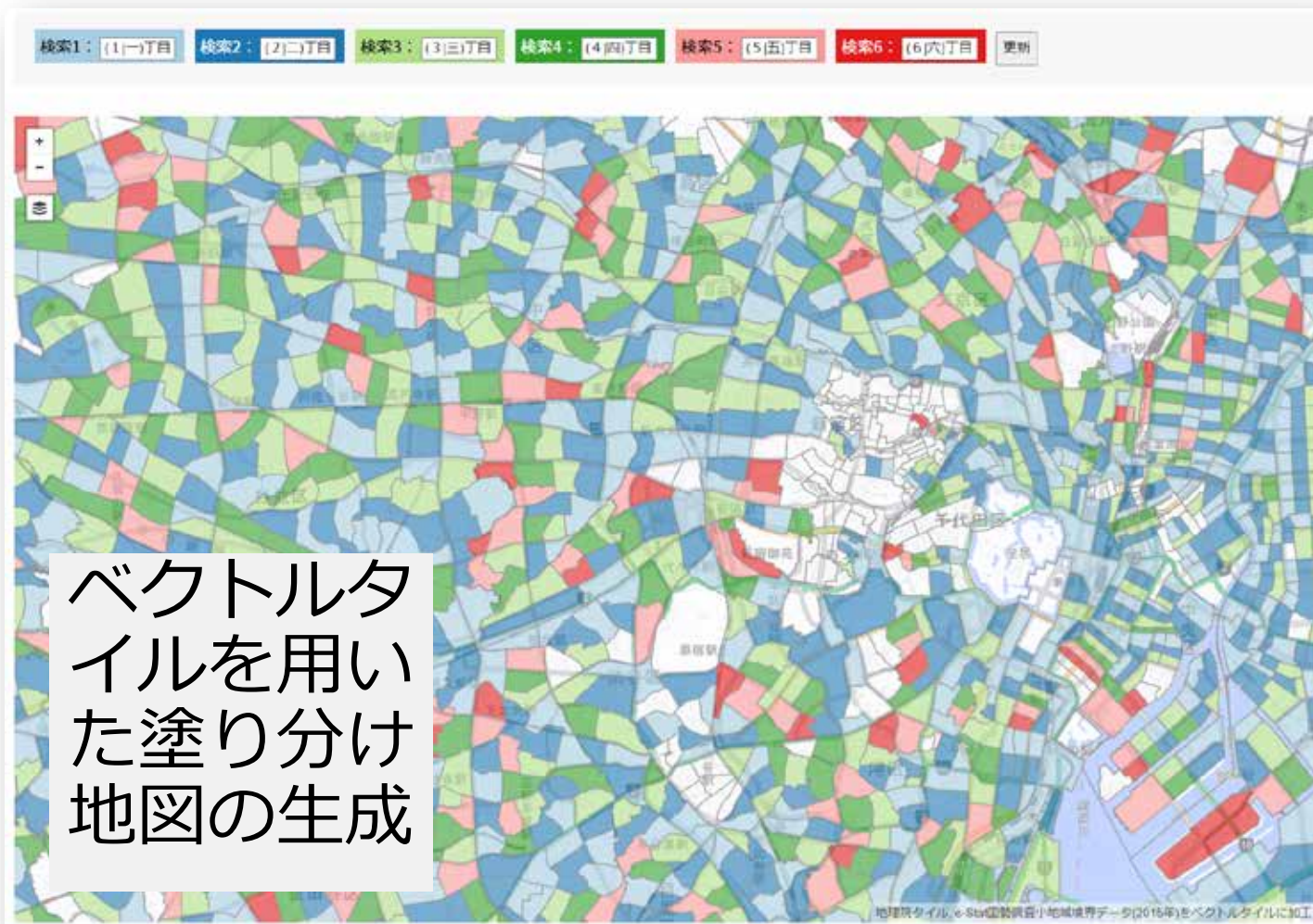
ズームレベル12
までは、市区町
村行政区域デー
タセットを表示。

ズームレベル13以降
は、町丁・字デー
タセットに切り替え。



地名ビジュアル検索

<http://geoshape.ex.nii.ac.jp/ka/visual-search/>



まとめ

1. エンティティを単位とする幾何データ共有サイト「**Geoshape**」を構築した。
2. GeoJSON、TopoJSON、ベクトルタイルなどの**ウェブマップ技術**を採用した。
3. 歴史的行政区域データセットβ版は、**現在～過去の市区町村データ**を統合した。
4. 国勢調査町丁・字等別境界データセットで**より細かい境界データ**を作成した。

今後の課題

1. より古い時代の地理データ（江戸・明治・大正時代等）に延長する。
2. 既存のデータセットの状況を把握し、連携や統合を検討する。
3. Geoshapeに加え、GeoNLP／GeoLODを、エンティティベースの地理情報ハブとして発展させる。
4. 防災等のアプリに展開していく。

参考情報

- 謝辞：本研究は、ROIS-DS公募型一般共同研究 ROIS-DS-JOINT 020RP2018および035RP2019（代表：村田 健史）の支援を受けました。
- 謝辞：データセットの構築には戸田智恵様の協力を得ました。
- Geoshapeリポジトリ：
 - <http://geoshape.ex.nii.ac.jp/>