

# 災害情報基盤構築に向けた地理情報リソースの整備

## Development of Geographic Information Resources toward Information Infrastructure for Disaster Management

榊 剛史 \*1,\*2    大谷 昭成 \*6    吉田 光男 \*2    鳥海 不二夫 \*2    篠田 孝祐 \*3    栗原 聡 \*3  
 Takeshi Sakaki    Akiyosi Otani    Mitsuo Yoshida    Fujio Toriumi    Kosuke Shinoda    Satoshi Kurihara  
 風間 一洋 \*4    野田 五十樹 \*5  
 Kazuhiro Kazama    Itsuki Noda

\*1株式会社ホットリンク  
Hottolink, Inc.

\*2東京大学  
The University of Tokyo

\*3電気通信大学  
The University of Electro-Communications

\*4和歌山大学  
Wakayama University

\*5産業技術総合研究所  
National Institute of Advanced Industrial Science and Technology

\*6慶應義塾大学  
Keio University

We are planning to develop information infrastructures for disaster management. As an example of those infrastructures, we develop a geographic information resource to assist digital volunteers who try to develop systems for disaster management. This system makes it possible to convert gps information and location name extracted from social media posts into administrative district information. Extraction of administrative district information is important for disaster management because support activities and rescue operations are performed by the administrative section in the case of disaster situation. Additionally, we share information about our systems, resources on our website [http://bit.ly/chidri\\_api](http://bit.ly/chidri_api).

### 1. はじめに

本近未来チャレンジセッション「異種協調型災害情報支援システム実現に向けた基盤技術の構築」は、今年で3年目を迎えることとなった。本セッションにおいては、1に示すように、災害発生時に、開発者が災害支援システムを容易に構築できるような情報基盤を構築することを最終目標としている。

これまで、本近未来チャレンジセッションにおいては、1年目は11件、2年目は10件の発表が行われ、3年目である今年は13件の発表が予定されている。その中で、情報伝播やシミュレーションの研究は数多く行われているものの、地理情報に関連する研究は少ない。これはGPS情報が付与された情報自体が少ない点、また地理情報処理に関する基盤が十分に整備されていないためと考える。

一方、災害発生時に、被災者を支援するための情報には地理情報が必要となる場合が多い。被災者の立場からすれば、通行可能経路や避難場所、必要な物資が提供されている場所などの情報が提供される場合、いずれもその位置が明らかになっていなければ、情報を活用することが困難である。また、災害支援者の立場からすれば、不足物資の支援、ボランティア活動を行う際に、やはりどこで何の物資不足しているのか、どこでボランティアが不足しているのかなど、同じ位置に関する情報が不可欠となる場合が多い。

実際、2011年3月11日に発生した東日本大地震においても、様々な情報が伝わらないことにより、災害支援活動に支障が生じた事例が多数挙げられる。例えば、各地での不足物資に関する情報が正しく伝達されなかったために需要を上回る支援物資が現地に送られ、その整理や補充に人や場所のリソースが割かれてしまうという事態が発生した\*1。逆に、地理情報を活用した災害支援事例も見受けられる。また、ホンダとパイオ

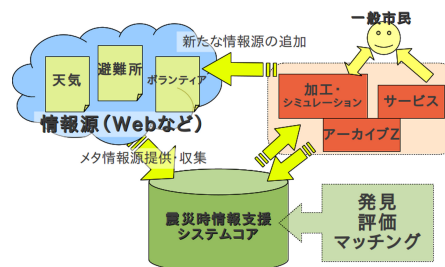


図 1:

ニアによる被災地の通行実績情報が Google MAP 上で提供されることで、通行可能な道をいち早く知ることができるようになり、被災者の避難行動や災害支援者が現地到着が効率化されたとされている\*2。それ以外にも不足物資情報、通行可能経路の情報、安否情報いずれの情報にしても、適切に活用するためには位置情報が不可欠である。

そこで本近未来チャレンジの3年目の成果として、災害情報支援に活用可能な地理情報基盤を整備し、それを公開する。本稿においては、地理情報基盤の持つ機能について紹介すると共に、その応用アプリケーション事例について紹介する。「災害情報支援システム」の実現を支援する枠組みを考案していく。なお、本研究で紹介した内容、開発システム、リソースに関する情報については、本近未来チャレンジの Web サイト上で共有する\*3

### 2. 既存の地理情報リソースの整備

本節では、現在利用可能な地理情報のリソースについてサーベイする。

連絡先: 榊 剛史, 株式会社ホットリンク, 東京都千代田区四番町 6 番 東急番町ビル

\*1 <http://communityarts.jp/wp-content/uploads/Asahi-Shinbun-11.06.11.pdf>

\*2 <http://response.jp/article/2011/03/15/153269.html>

\*3 異種協調型災害情報支援システム実現に向けた基盤技術の構築 (CHIDRI)  
<https://sites.google.com/site/crepchidri/home>

## 2.1 空間情報科学研究センター

空間的な位置や領域を明示した自然・社会・経済・文化的な属性データを系統的に、構築、管理、分析、統合、伝達するための汎用的な方法とその応用を考える「空間情報科学」を研究を目的とした組織である。様々な地理情報リソースを提供しているが、下記ではテキストにおける地名情報処理に活用可能なリソースのみを紹介する。

- アドレスマッチングサービス [空間 15a]
- シンプルジオコーディング実験 [空間 15b]
- ジオコーダ DAMS [空間 15c]

## 2.2 GeoNLP プロジェクト

自然言語文にジオタギングするための地名情報処理システムを構築するためのプロジェクトである。国立情報学研究所により運営・提供されており、下記3つのサイトが提供されている。

**GeoNLP** 地名に関するオープンデータを共有するサイト [国立 15b]

**GeoNLP 地名ウェブサービス** 地名辞書を使って自然言語文を解析するサイト [国立 15c]

**GeoLOD** 地名に関する Linked Open Data (LOD) を提供するサイト [国立 15a]

特に GeoNLP のウェブサイトにおいては、多様な地名辞書が提供されているだけでなく、ユーザが地名辞書を共有するためのシステム及びそのための地名辞書フォーマットが提供されており、地名辞書を検索・共有する上では有意義なサイトとなっている。

## 2.3 その他活用可能な地理情報データ

本節では、原稿執筆時点で利用可能な日本国内の地理に関する地理情報リソースをまとめた。

### 2.3.1 国立・政府機関が配布するリソース

- 国土交通省国土政策局 GIS データ閲覧・ダウンロードサービス [国土 15a]
- 電子国土基本図（地名情報） [国土 15b]
- 郵便番号データ [郵便 15]

### 2.3.2 その他企業・個人が配布するリソース

- 日本行政区画便覧データファイル [日本 15]
- GSK 地名施設名辞書第2版 [言語 15]
- 10万語地名辞書 [ato 11]
- 山岳名辞書 [ato 11]
- 統一地名辞書 [ato 11]

### 2.3.3 参考となるリソース

その他、配布を目的としたものではないが、参考となるリソースについて紹介する。

- はてなキーワード 地理カテゴリ <sup>\*4</sup>
- Wikipedia Category:日本の地理 <sup>\*5</sup>

\*4 <http://d.hatena.ne.jp/hotkeyword?cname=geography>

\*5 <http://ja.wikipedia.org/wiki/Category:日本の地理>

## 3. 地理情報基盤に求められる機能

本節では、災害支援のための地理情報基盤に求められる機能について整理する。

### 3.1 位置情報の変換

地理情報処理を行う上で、地名と緯度経度情報の相互変換は必須となる。たとえば、ユーザのソーシャルメディアへの投稿から GPS による緯度経度情報が取得できたとしても、投稿者以外がその情報を活用するためには、人が理解可能な地名に変換する必要がある。また、投稿者自身も、自分の現在位置を避難などに活用するためには、緯度経度情報から行政区分を把握したり、近くのランドマークを検索したりする必要がある。

このように数値情報である緯度経度情報とその数値の指す地名、ランドマーク、行政区分などは相互変換する必要がある。そこで、構築する情報基盤では下記のような位置情報変換機能を実装する。なお、地名から緯度経度情報への変換はジオコーディング、緯度経度情報から地名への変換を逆ジオコーディングと呼ばれる。

- 緯度経度 → 行政区域変換（都道府県単位、区市町村単位、大字・小字単位）
- 行政区域 → 代表地点の緯度経度変換
- 地名（建物・ランドマーク名） → 代表地点の緯度経度変換

### 3.2 位置情報の生成

ある情報に地理情報処理を行う際は、その情報に緯度経度情報が付与されていることが必要となる。しかし、現実には、すべての情報に緯度経度情報が付与されているわけではない。もちろん、位置と紐付けることができない情報も多数あるが、位置としての手がかり情報を持つ情報は一定量存在する。たとえば、ソーシャルメディア上の投稿であれば、GPS 情報が付与されていなくても、地名を含む投稿は多数存在する。このような位置情報の手がかり情報から位置情報が生成できれば、それらについて、地理情報処理を適用することが可能になる。

そこで、本構築基盤では、位置情報生成機能を実装する。

- 投稿内容からの位置情報生成
- ユーザ行動からの位置情報生成

## 4. 機能実現にむけて

### 4.1 位置情報の変換

ジオコーディング、逆ジオコーディングを行うためには、地名辞書およびそれに対応する位置情報データの整備が必須となる。本研究では、まず前述した各地名辞書を収集・変換し、統一的な地名辞書の整備をおこなった。今回の地名辞書整備に用いたのは下記の通りである。

- GeoNLP にアップロードされている辞書
- 国土政策局による位置情報データ
- 郵便番号データ
- GSK 施設地名辞書第2版
- 10万語地名辞書

- 山岳地名辞書
- はてなキーワードの地理カテゴリに含まれるワードのうち位置情報が付与されているもの、
- Wikipedia のエントリにおいて「日本の地理」カテゴリに属し、かつ位置情報が付与されているもの

これらの地名辞書から地名および緯度経度情報を抽出した後、地名と緯度経度情報をペアとしてデータベースの格納した。

## 4.2 位置情報の生成

位置情報の生成手法は、投稿への位置情報生成とユーザ行動からの位置情報生成に分けられる。

### 4.2.1 投稿内容からの位置情報生成

投稿内容からの位置情報生成については、投稿から地名に当たる語を抽出し、その地名を位置情報に変換することで実現できると考えられる。よって、前述の様に辞書を整備することで実現できると考えられる。しかし、特にソーシャルメディアのように口語的な投稿が数多くなされる場合、ユーザが必ずしも地名と同じ表記でその地名を投稿するとは限らない。そこで、より砕けた地名表現を判定し、それを既存の地名辞書と紐付ける技術が必要となる。文書からの地名抽出については、固有表現抽出を持ちいたアプローチが一般的であるが、それだけでは抽出した地名に位置情報を紐付けることが困難であるため、固有表現抽出のみでは不十分である。

本システムでは、口語的な地名表現に対しては、Neural Network 言語モデルを用いた地名判定手法により、各地名の口コミ表現を獲得する手法を用いる [森國 15a]。その地名辞書を元にと、投稿内容から位置情報を推定する手法としては、森國らの手法を用いる予定である [森國 15b]。森國らの手法では、投稿位置推定においてノイズとなる単語を除く TF-IAF フィルタという手法を提案するとともに、Cheng らの提案した単語出現頻度のスムージング方法を用いることで、従来手法よりも高い精度で投稿内容からの位置情報推定を実現している [Cheng 10]。

### 4.2.2 ユーザ行動からの位置情報生成

ユーザ行動からの位置情報生成については、各ユーザの普段の行動からユーザの居住地を推定し、それを位置情報の変換することで実現できると考えられる。ソーシャルメディアの普及に伴い、近年ではユーザの一連の投稿から、ユーザの居住地を推定する手法が提案されている。

ユーザの居住地推定手法としては、Cheng 手法がよく用いられているが、ユーザの一連の投稿に対して、ツイート単位で位置情報を推定した後に、それらを集計することで、居住地推定を実現している [Cheng 10]。そのため、各ツイートに対する位置情報の推定は前述の森國らの手法を適用することができると考えられる。そこで本システムでは、Cheng らのアプローチを採用しながら、各ツイートの位置情報推定には森國らの手法を用いることで、ユーザの居住地推定を実現する。

## 5. 位置情報基盤の実装および提供方法

### 5.1 実装方法

位置情報基盤システムを構築するアプローチはいくつか考えられる。位置情報を扱える DB は限られるため、DB 決定することで自ずと選択肢が限られてくる。

本システムにおいては、Elastic Search (ES) を用いて位置情報基盤を実装することとする。Elastic Search を用いる理由は下記の通りである。

- 位置情報を蓄積し、検索する機能が標準で実装されている
- JSON を用いた REST API (後述) が標準で実装されている。
- 大規模分散フレームワークであるため、大規模化が容易である
- 活用事例が豊富であり、活用・運用のための知見がウェブ上に豊富に蓄積している
- 活用事例が豊富であるため、開発者が取り組みやすい

REST (Representational State Transfer) API とは HTTP をベースとしたプログラム用のインターフェースであり、Twitter や Facebook, Amazon Web Services など様々なウェブシステムが REST 形式で API を公開している。アクセスに特殊なクライアントを必要としないためアプリケーションの開発が容易になる、API が規律化されるため設計がしやすくなる、などのメリットがあげられる。ES では、標準で REST API が準備されているため、API を設計・実装することなく、容易に他のシステムと連携させることが可能である。

### 5.2 提供方法

本システムの提供方法は2つである。API 方式と AMI 方式である。

#### 5.2.1 API 方式

API 方式とは、先に説明したように REST 形式の API により、機能を提供する方式である。実際には、ES の API により提供が可能であるため、ES に準拠した方式で API のクエリを作成することができれば、本システムの機能およびデータを活用することができる。そのため、開発者は非常に容易に本システムを活用することができる。その反面、事前に登録が必要である点、ウェブを介するために大量のデータを高速に処理することが困難である点、本プロジェクトのサーバに障害は発生した際に使用不能となる点がデメリットとしてあげられる。

#### 5.2.2 AMI 方式

AMI とは Amazon Machine Image の略であり、Amazon Web Service 上で提供される Elastic Compute Cloud で仮想サーバ (インスタンス) に起動に必要な情報を提供するファイルである。本方式では、地理情報基盤を起動するために必要なツールおよびデータがすべて入った状態の AMI を提供する。AWS では AMI からインスタンスを生成することで、手間のかかる環境構築、必要なソフトウェアのインストールをすることなく、すぐに本システム基盤を起動することができる。各開発者の手で本システム基盤を運用することが可能になるため、大量データに対する高速な処理が可能、事前登録が不要、サーバー障害時は別のサーバをすぐに起動できるため耐災害性・障害性に優れる、などをメリットがあげられる。デメリットとしては、開発者に AWS の運用知識が必要である点、コストがかかる点などがあげられる。

上記2つの提供方式を採用することで、より柔軟にシステム開発が可能な地理情報基盤を提供することを目指す。

### 5.3 位置情報を用いた応用技術

位置情報を用いた応用技術、すなわちアプリケーションとしては、様々なものが考えられる。

まず自分の現在位置から近いランドマークや避難所、物資の配給場所などをといった場所に紐付いた検索は本システムに避難所、ランドマーク、配給場所の情報を登録することができれば、ES の検索機能で容易に実現することができる。

まだより高度な研究をあげると、村上らの研究として、東日本大震災直後のツイートデータから不足物資を推定する研究があるが、これに位置情報を用いることで、地域ごとに不足している物資を推定することができるようになると考えられる [Murakami 12].

また、谷らの研究で、Twitter 上から交通経路を抽出する研究がある [谷直 15]. この研究は、ジオタグ付きツイートを用いているため、位置情報の精度が高い反面、位置情報を持つツイートが少ない問題点がある。本システムを用いてツイート内容から位置情報を生成することができれば、より広い範囲について位置情報を取得し、交通経路を抽出できる可能性がある。

## 6. 終わりに

本稿では、開発者が災害支援システムを構築する上で必要となる可能性が高いシステムとして、地理情報基盤システムの提案、構築および共有を行った。

まず、既存の地理情報リソースを整理し、現時点で利用可能なリソースをまとめた。次に地理情報基盤に必要と思われる機能を整理したのち、それらの機能の実現手法について述べた。最後に、開発者がより使いやすい形での地理情報基盤の実装方法および提供方法について説明した。

このような地理情報基盤システムは、有事の際に短期間で使うのは困難であると考えられるため、今後は、平時から本システムを運営していくものとする。次年度以降、本システムを活用した研究およびシステムが増えていくことを期待していきたい。

## 7. 謝辞

本研究の執筆にあたり、位置情報の有用性や活用方法についてアドバイスを与えてくれた筑波大学の原様に感謝いたします。また、本研究は課題設定による先導的人文・社会科学研究推進事業、「リスク社会におけるメディアの発達と公共性の構造転換～ネットワーク・モデルの比較行動学に基づく理論・実証・シミュレーション分析」の助成を受けました。

## 参考文献

- [ato 11] 書き屋のための変換辞書 for ATOK, <http://homepage3.nifty.com/t-weekly/dictionaries4.html> (2011)
- [Cheng 10] Cheng, Z., Caverlee, J., and Lee, K.: You Are Where You Tweet : A Content-based Approach to Geo-Locating Twitter Users, in *Proceedings of the 19th ACM International Conference on Information and Knowledge Management, CIKM '10*, pp. 759–768, ACM Press (2010)
- [Murakami 12] Murakami, A. and Nasukawa, T.: Tweeting about the tsunami?: mining twitter for information on the tohoku earthquake and tsunami, in *Proceedings of the WWW 2012 Companion, WWW '12 Companion*, pp. 709–710, New York, NY, USA (2012), ACM
- [空間 15a] 空間情報科学研究センター: アドレスマッチングサービス, [http://newspat.csis.u-tokyo.ac.jp/geocode/modules/addmatch/index.php?content\\_id=1](http://newspat.csis.u-tokyo.ac.jp/geocode/modules/addmatch/index.php?content_id=1) (2015)
- [空間 15b] 空間情報科学研究センター: シンプルジオコーディング実験, [http://newspat.csis.u-tokyo.ac.jp/geocode/modules/geocode/index.php?content\\_id=1](http://newspat.csis.u-tokyo.ac.jp/geocode/modules/geocode/index.php?content_id=1) (2015)
- [空間 15c] 空間情報科学研究センター: ジオコーダ DAMS, [http://newspat.csis.u-tokyo.ac.jp/geocode/modules/dams/index.php?content\\_id=1](http://newspat.csis.u-tokyo.ac.jp/geocode/modules/dams/index.php?content_id=1) (2015)
- [言語 15] 言語資源協会: GSK 地名施設名辞書第 2 版, <http://www.gsk.or.jp/catalog/gsk2012-c/> (2015)
- [国土 15a] 国土交通省国土政策局: GIS データ閲覧・ダウンロードサービス, <http://nlftp.mlit.go.jp/index.html> (2015)
- [国土 15b] 国土交通省国土地理院: 電子国土基本図 (地名情報), <http://www.gsi.go.jp/kihonjohochousa/chimeijoho.html> (2015)
- [国立 15a] 国立情報学研究所: GeoLOD, <http://geolod.ex.nii.ac.jp/> (2015)
- [国立 15b] 国立情報学研究所: GeoNLP, <https://geonlp.ex.nii.ac.jp/> (2015)
- [国立 15c] 国立情報学研究所: GeoNLP 地名ウェブサービス, <https://dias.ex.nii.ac.jp/geonlp> (2015)
- [森國 15a] 森國泰平, 吉田光男, 岡部正幸, 梅村恭司: ツイート投稿位置推定のためのノイズとなる単語の除去手法, 第 7 回データ工学と情報マネジメントに関するフォーラム (2015)
- [森國 15b] 森國泰平, 吉田光男, 岡部正幸, 梅村恭司: ツイート投稿位置推定のためのノイズとなる単語の除去手法, 第 7 回データ工学と情報マネジメントに関するフォーラム (2015)
- [谷直 15] 谷直樹, 風間一洋, 榎剛史, 吉田光男: ジオタグ付きツイートをを用いた交通路の抽出, 第 7 回データ工学と情報マネジメントに関するフォーラム (2015)
- [日本 15] 日本加除出版: 日本行政区画便覧データファイル, <http://www.kajo.co.jp/denshi/denshi01.php> (2015)
- [郵便 15] 郵便局: 郵便番号データ, <http://www.post.japanpost.jp/zipcode/download.html> (2015)