

# 移動データの構造化に向けた入力支援手法の開発と 移動情報の流通可能性に関する研究

熊谷潤・金杉洋・柴崎亮介・山口琢・小林龍生

Study on Development of Input Method for Structuring of Human Flow Data and  
Possibility of Reusing and Distribution of Human Flow Information

Jun KUMAGAI, Hiroshi KANASUGI, Ryousuke SHIBASAKI,  
Taku YAMAGUCHI, Tatsuo KOBAYASHI

**Abstract:** Now days, CGM like BLOG, SNS is spreading out in all generations, and the amount of a data which contained human flow data like diary or travel is increasing. Enterprises or the government tried to extract valuable information for them from these data. We study on how to structure easy input system for everyone, and the possibility of reuse and distribution of this human flow information. As for the input method, we developed a new input system which contains both usual way of text input and structuring extendibility of XML. Also through a web application that handles the human flow information, we study on the possibility of reuse and distribution of the human flow information.

**Keywords:** 移動情報(Human Flow Information), 構造化(Structuring), 入力支援手法(Input Method)

## 1. はじめに

現在ブログやSNS, 口コミサイトなどのCGMが普及し, 行動データに関しても, ブログ日記や旅行記をはじめ, スケジューラーやメールデータなどの個人データ, また旅行会社のツアーデータ, 地方自治体の観光データ, パーソントリップデータなど我々の環境において多く散在している。

---

熊谷, 金杉, 柴崎 : 〒277-8568 千葉県柏市柏の葉5-1-5  
東京大学 空間情報科学研究センター  
山口, 小林 : 株式会社ジャストシステム

これらの大量の行動データの中から必要な行動情報を拾い出し活用したいというニーズも高まっている。

こうした状況の中, ブログや WEB サイトなどのテキストデータから特定の情報を抽出しようとする研究がなされている。またブログや文書データから場所情報を抽出し実空間との関連付けを行う研究(相良・有川,2004) (佐藤・西山・砂山,2007)がなされている。さらに経路情報を抽出するという研究(群ほか,2006)がなされはじめて

おり、実世界の人間の行動情報への関心も高まってきた。しかしながら、これらの研究は主にテキストマイニングによる手法であり、項目別に抽出手法を研究する必要があることや、抽出精度の限界などの課題があげられる。

一方で XML によって、データをタグで囲み構造化する試みがなされている。XML の入力手法を見てみると、XML の直書きエディタがあげられる。しかし、一般人にとって XML を手書きするのは、かなり困難なことと言える。移動情報を扱った既存研究の入力手法についてみてみると、交通行動調査票の手書きのものから、WEBTFP(大藤ほか,2005)や WEB ダイアリーのような WEB 項目入力を使用したものがある。しかし、ユーザにとって数ある項目を入力していくのは面白みがなく面倒であると考えられる。

## 2. 研究目的

本研究では、散在する移動データから価値ある移動情報を構築し利用するために、慣れ親しんだテキスト入力の簡易性と XML の構造化拡張性を兼ね備えた新しい入力支援手法の開発を行う。また構造化された移動情報を共有可能な応用アプリケーションを通じて、移動情報の構造単位での再利用・流通可能性に関する研究を行う。

## 3. 入力支援手法

### 3.1. 移動データの構造化

移動データを構造化する上で、場所に関する情

報は必須であるといえる。次に、移動情報を現実の時間軸に即したものとするため、日時情報が必要である。最後に、抽出するテキストデータについて、どこまでの範囲が該当の場所や日時に関する記述なのかの判定や場所情報と日時情報の関連付けを行うためにイベント区切りの情報が必要である。本研究では、場所、日時（開始日時、終了日時）、イベント区切り、この 3 つの構造化項目を元にテキストデータを構造化する入力支援手法について研究する。

### 3.2. 入力支援手法

入力支援手法に関して、文章作成において必ず使用されている仮名漢字変換機能に着目した。また外部から場所や日時に関する情報を取り込むために、ユーザ辞書として入力支援システムと外部サーバとの通信による連携に着目した。本研究では、この 2 つの着目点をもとに、入力支援システムを開発した。これにより、テキスト入力時と同様に変換を行い、誰もが簡単に関連情報を含んだタグを入力することができる。

#### 3.2.1. フォーマット

入力支援システムで入力されるフォーマットは、microformats を採用し拡張した。

- ・場所タグ : <span class="location">～</span>
- ・開始時間タグ : <abbr class="dtstart">～</abbr>
- ・イベントタグ : <div class="vevent">～</div>

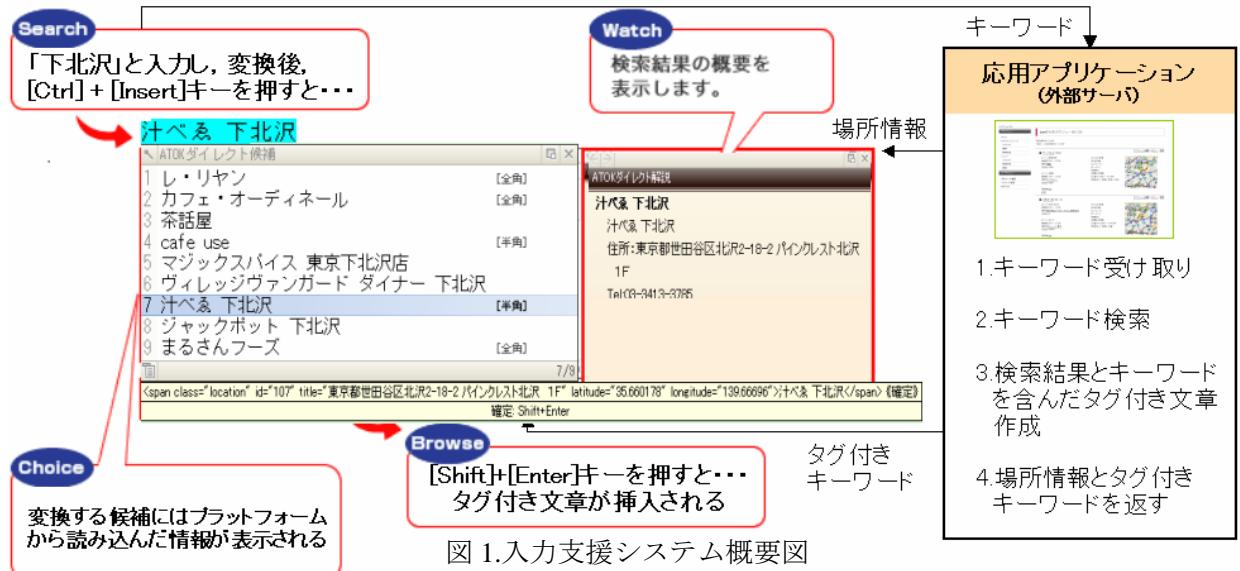


図 1. 入力支援システム概要図

### 3.2.2. 変換候補値

タグ内に挿入されるキーワードに関する関連情報に関しては、外部サーバとの通信により、関連キーワードについて検索された結果情報を取得し、タグ内に埋め込む。

### 3.3. 場所情報入力支援

ここでは、「下北沢」と入力した場合を例に、入力支援システムの挙動について説明する。

- 「下北沢」と入力を受け取った入力支援システムは、外部サーバのアプリケーションに API を通じて「下北沢」で検索を行う。
- アプリケーションは、「下北沢」に関するショッピングや施設などの場所に関する検索結果を入力支援システム側に API 配信を行う。
- 入力支援システムは、アプリケーションから受け取った XML の結果を解析し、場所タグ <span class="location"> 内に返値を挿入し、入力候補値として表示する。
- ユーザは「下北沢」に関する返値候補から自分の入力したい候補値を選択入力する。

### 3.4. 日時情報入力支援

ここでは、「明日の午後 3 時」を例に入力支援システムの挙動を説明する。

- 「明日の午後 3 時」と入力を受け取った入力支援システムは、「明日」と「午後 3 時」の語句を抽出し、現在日時から「明日」「午後 3 時」を計算し正しい日時の値を求める。
- 開始日時タグ dtstart の title 属性値に返値を挿入し、入力候補値として表示する。
- ユーザは、返された候補値から自分の入力したい日時（開始か終了）を選択入力する。

### 3.5. イベントタグ入力

イベントタグの入力時の入力支援システムの挙動は以下の通りである。

- 「イベント」と入力を受け取った入力支援システムは、開始タグ 

<div class="vevent">

 および終了タグ 

</div>

 を値として返す。

- ユーザは、イベント開始もしくは終了にあわせ該当のイベントタグを選択入力する。

### 3.6. 入力支援手法による移動データ構造化例

入力支援システムによって、構造化された移動データの文章データの例を以下に示す。

#### 1. 結果表示

今日 7 時ころから、六本木で映画をみてきた。

#### 2. 結果ソース

```
<div class="vevent"><abbr class="dtstart" title="2008/08/16 19:00">今日 7 時ころから、</abbr><span class="location" id="2" name="TOHO シネマズ六本木" latitude="35.660175" longitude="139.728958" title="東京都港区六本木6-10-2 六本木ヒルズけやき坂コンプレックス内">六本木</span>で映画をみてきた。</div>
```

### 3.7. 入力支援手法によるメリット

本入力支援システムにより、文章を書く際に、タグ変換し移動データの構造化入力が可能になる。また文字を入力できる場合どこでも使用可能である。様々な専門データベースと連携することにより、関連情報をタグ内に挿入し入力できる。

## 4. 応用アプリケーション「モスケ」

本研究では移動情報の応用例として、移動情報を構造単位で投稿、収集、表示、検索、出力可能な WEB アプリケーション「モスケ」を開発した。ブログ記事上の移動情報タグを解析し、モデルレスケジュールとして変換保存する機能を実装している。また蓄積されたデータは、API 配信され、入力支援システムの地名辞書として活用される。

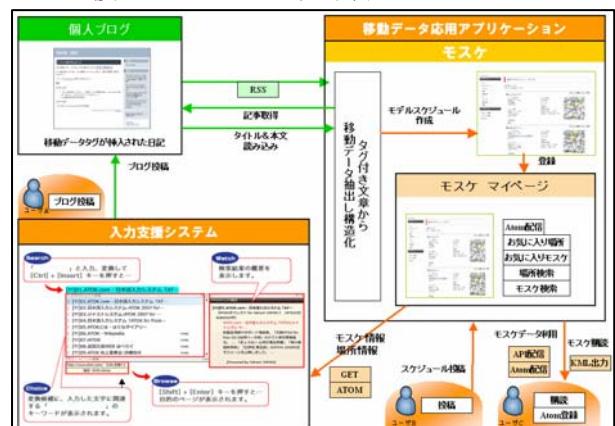


图 2. アプリケーションシステム概要図

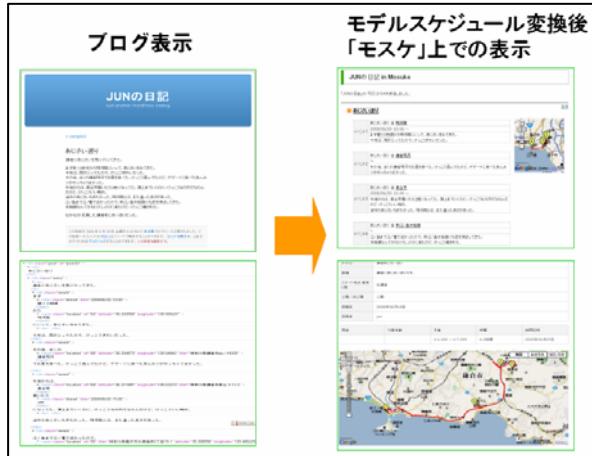


図3.ブログからの移動情報抽出例

#### 4.1. 「モスケ」によるメリット

「モスケ」により、単なる文章データから経路化された移動情報の地図表示や、スケジューラー表示など詳細な視覚化が可能になる。また共有された移動情報を場所などの構造単位で検索、参照、転用が可能になり、生活に活かすことができる。さらに「モスケ」に蓄積された場所情報を、入力支援システムの地名辞書としての再利用、ブログや外部アプリなどへ配信することができる。また行動マーケティングやパーソントリップなど動線解析などへの応用も考えられる。

#### 5.まとめ

本研究では、散在する移動データを構造化し価値ある情報として利用するため、我々が慣れ親しんだテキスト入力の簡易さと XML の構造化拡張性を兼ね備えた新しい入力支援手法を示した。さらに構造化された移動情報を構造単位での検索や再利用・流通を可能にする応用アプリケーション「モスケ」による活用例やメリットを示した。応用アプリケーション上に移動情報が蓄積されることにより、入力支援システムの地名辞書として再利用することが可能になる。これによりさらに移動情報の構造化入力が利用しやすいものとなり移動情報が流通しやすくなる。このように本研究の入力支援システム、応用アプリケーションによって移動データを構造化し共有することによって、移動情報の再利用のスパイラルが形成さ

れる。以上のことから、誰もが移動データを簡単に構造化でき、価値ある情報として再利用・流通できる仕組みが整ったと考えられる。

今後、入力支援システムの有効性の検証として、マニュアルと入力支援システム使用時での移動情報にかかる所要時間や容易さを計測比較、またユーザが意図したとおり、ブログ記事から移動データを抽出構造化できたかどうかの抽出精度、ユーザの反応などを実験評価する予定である。

また移動情報の再利用・流通可能性について、エリア限定で観光データ、ツアーデータ、ブログデータ、スケジュールデータなどを様々な移動データを構造化し蓄積された移動情報を構造化の特性を活かしてどのように利用できるか、またどのようにユーザに利用されるのかを評価検証していく予定である。

#### 参考文献

- [1]. WEB ダイアリー,  
"http://www.probe-data.jp/demo/diary.html"(2008/06/27 アクセス)
- [2]. 大藤武彦,松場圭一,井上英樹,松村暢彦 (2005) 「WEB を活用したトラベル・フィードバック・プログラムの多様な事業所への適用」土木計画学研究・講演集 CD-ROM Vol.31
- [3]. 郡宏志, 服部峻, 手塚太郎, 田島敬史, 田中克己 (2006) ”ブログからのビジターの代表的行動経路とコンテキストの抽出”, 情報処理学会研究報告,p35-p42
- [4]. 相良毅, 有川正俊 (2004) ”ジオベースによる Web からの空間コンテンツ獲得” 電子情報通信学会データ工学ワークショップ,I-11-01
- [5]. 佐藤圭太, 西原陽子, 砂山渡 (2007) ”ブログ記事からのイベント文抽出によるシーンの生成”, 人工知能学会