

平成28年度研究開発成果概要書

採択番号：178B05

課題名：ソーシャル・ビッグデータ利活用・基盤技術の研究開発

個別課題名：課題B 新たなソーシャル・ビッグデータ利活用・基盤技術の研究開発

副題：ソーシャル・ビッグデータ駆動の観光・防災政策決定支援基盤の研究開発

(1) 研究開発の目的

本研究開発は、Web データ収集・利活用基盤の構築、社会データを活用した科学的分析に基づく合理的な公共サービスの提供、社会データ基盤の持続的運用可能なビジネスモデル開発と、これらの社会実装を目的とする。

観光など地域経済活性化や防災・減災政策などの問題解決の難しさは、部分的でしかも不完全な情報やデータに基づいて、リスクや利益を推定し、主観的判断によって、意思決定を行わなくてはならないことにある。そこで、社会経済分野の多種多量な情報やデータを収集し、科学的分析手法に基づいてデータを解析し、国や自治体の政策決定や企業などの経営の意思決定を支援するICTシステムとサービス開発基盤を構築する。

具体的には、情報空間にアップロードされる様々なデータを収集する社会データ収集基盤を構築する。また、自治体がつもつオープンデータや、Wi-Fi アクセスポイントのログデータを連携させ、俯瞰的に眺めるツールや、各種データの連携・分析によるサービスの合成や政策・意思決定支援を行うソーシャル・ビッグデータ駆動のデータ中心政策決定支援基盤を実現する。

本研究開発で扱う社会データは、国勢調査などの公的統計データ（これは、e-Gov データ、オープンデータと呼ばれる）、宿泊施設、観光施設、公共交通のWeb 予約データや賃貸不動産データ（“BOOK-log” と称す）、Wi-Fi アクセスポイントのアクセスログ、天気や台風、地震などの気象データ、地凶データ、食データ(FOOD-log)からなる。これらデータを中心とした科学的分析を行い、人やモノを制御する情報やサービスを合成し、迅速かつタイムリーにフィードバックする技術的・社会的仕組み、および、主に観光客を対象とした行動予測を実現する。特に、本研究開発・実用化では、宿泊施設、観光施設（ゴルフ場）、高速バスのWeb 予約データ、賃貸不動産データであるBOOK-log およびWi-Fi アクセスログを対象とし、データの収集・蓄積・共有を行うための高度社会データ基盤を構築し、不良設定逆問題解法としてのデータ統合方法を確立し、人やモノを制御するICT 情報サービスを合成して、迅速かつタイムリーにフィードバックするデータ駆動政策決定支援システムを研究開発・実用化する。

(2) 研究開発期間

平成26年度から平成29年度（4年間）

(3) 実施機関

大学共同利用機関法人情報・システム研究機構 国立情報学研究所

（実施責任者 研究主幹・教授 曾根原登）＜代表研究者＞

大学共同利用機関法人情報・システム研究機構 統計数理研究所

（実施責任者 研究主幹・教授 山下智志）

国立大学法人山梨大学

(28-1)

(実施責任者 准教授 渡辺喜道)  
国立大学法人長崎大学  
(実施責任者 教授 小林透)

(4) 研究開発予算（契約額）

総額 120 百万円（平成 28 年度 25 百万円）  
※百万円未満切り上げ

(5) 研究開発項目と担当

**研究項目 1**： 山梨県・ハケ岳観光圏を対象とした観光政策・意思決定支援基盤の研究開発（山梨大学）

1. ペンションなど宿泊施設およびゴルフ場の BOOK-log の収集基盤の構築および収集データを利用した稼働率推定手法の確立
2. 山梨県の路線バスを対象とした人流推定手法および運行計画の最適化手法の確立
3. 路線バスと宿泊施設の連携による観光客分布予報の確立
4. ハケ岳観光圏を対象とした実態調査の実施

**研究項目 2**： 長崎県を対象とした観光政策・意思決定支援基盤の研究開発（長崎大学）

1. Wi-Fi アクセスログおよび GPS データの有用性の実証とリアルタイム観光客分布の可視化システムの構築
2. Wi-Fi アクセスログおよび GPS データ、統計データの連携による観光客数および分布推定手法の開発
3. 観光船の入港データおよび船内の Wi-Fi アクセスログに基づく外国人観光客を対象とした観光客予報の確立
4. 観光客が発信した情報を収集・追加し、観光地の最新情報を発信する“自己拡張型オープンデータプラットフォーム”の確立

**研究項目 3**： 民泊を想定した賃貸不動産の占有状態の推定モデルの研究開発（統計数理研究所）

1. 賃貸不動産情報の収集の効率化
2. 賃貸不動産の占有状態遷移モデルの構築のための実地調査
3. 占有状態の促進のための要素の解明

**研究項目 4**： 観光政策・意思決定支援基盤の共通プラットフォームの開発およびモジュール化に関する研究開発（国立情報学研究所）

1. Web/SNS 空間のデータ収集・可視化システムのモジュール化および API 標準化
2. 高速バスを対象とした Web/SNS データに基づくサービス合成の実証実験
3. 観光政策・意思決定支援の共通プラットフォームの構築

## (6) これまで得られた成果（特許出願や論文発表等）

		累計（件）	当該年度（件）
特許出願	国内出願	0	0
	外国出願	0	0
外部発表	研究論文	11	1
	その他研究発表	166	39
	プレスリリース・報道	11	11
	展示会	5	1
	標準化提案	0	0

## (7) 具体的な実施内容と成果

**研究項目1**：山梨県・八ヶ岳観光圏を対象とした観光政策・意思決定支援基盤の研究開発（山梨大学）

山梨県の路線バスを対象とした人流推定手法の基礎研究として、路線バス運行状況表示システム（バスコンシェルジュ）とのデータの連携を行い、遅延状況の可視化など運行サービス改善に資する基礎分析を行った。15秒間隔で蓄積されるデータをもとに運行の遅延に関する分析を行い、ボトルネックとなる箇所の検出や周期性に関する検証を行った。また八ヶ岳エリアでの観光客の動態を把握するために、主要な観光スポットでのセンシング情報をもとに人流を可視化するシステムの検討を行った。

**研究項目2**：長崎県を対象とした観光政策・意思決定支援基盤の研究開発（長崎大学）

Wi-Fi アクセスログ等による観光客動態の推定を行った場合の問題点として、地方においてはそのサンプル数の偏りから現実の状態を上手く反映していないことが挙げられる。これに対し、得られたログデータから地方の現状を推定する手法について公的統計データを活用した推定モデルの構築を行った。ログデータの生成に関するバイアスならびに統計データによる限定的な情報量により継続的にデータとモデルの検証を行う。

また自己拡張型オープンデータプラットフォームとして、ロケーションツーリズムに適用した「ながさきロケナビ」を開発、他観光コンテンツへのポータビリティ確保のため、構造をMVCモデル化することで、新規コンテンツの適用における効果を検証し、保守性および拡張性が高いことを確認した。

これまでに蓄積される研究成果の実用化に向けて、長崎県の自治体クラウド上で稼働する可視化システムとの宿泊施設稼働率・Wi-Fi アクセスログによる観光客検出データの連携に関する基礎開発を実施した。

**研究項目3**：民泊を想定した賃貸不動産の空占状態の推定モデルの研究開発（統計数理研究所）

Webビッグデータとして宿泊施設予約状況のデータと不動産鑑定士によるデータのそれぞれのデータ取得とモデル化を行った。不動産鑑定士によるデータを用いて空占モデルを構築し、空占の遷移を物件条件からモデル化することができる可能性があることを示した。またWebビッグデータに対する構造化手法（前処理）を開発し、それらのデータの基本的な性質の調査と簡易モデルとしての空占のモデル化を実施した。これによりWebビッグデータが不動産データをある水準で近似する基本的な統計であることを確認した。

**研究項目4**：観光政策・意思決定支援基盤の共通プラットフォームの開発およびモジュール化に関する研究開発（国立情報学研究所）

Web/SNS空間のデータ収集・可視化システムのモジュール化およびAPI標準化として、モジュール化ならびにTTPP（TrackId-TimeWindow-Place-Pair）を用いた状態の差分ベクトルによるAPI標準化を行った。これによりカラム方向には2倍の領域を必要とするものの、n件

(28-1)

のエントリーに対して多項式時間での比較が必要であったものが、線形時間での比較で情報の取得が可能となった。これにより各種予約サービスから取得される Web ビッグデータを TTPP の標準化により統一的に扱えるようになった。他のデータとの掛け合わせによるサービス合成のための、素データレベルでの検証を行った。

また、観光政策・意思決定支援の共通プラットフォームの構築として、クローラー・格納・集計に関するモジュールの整備を行った。研究項目4の1による API を Node-RED を用いて呼び出すことで、集計サービスの作成ができるよう整備を行った。また一部のモジュールを長崎大学において運用することにより、長崎クラウド上での可視化のための整備を行った。