

世界メッシュコードに基づく各種データの可視化

Visualization platform for various data based on world grid square codes

佐藤 彰洋[†] 西村 正貴[‡] 椿 広計[‡]

Aki-Hiro Sato Shoki Nishimura Hiroe Tsubaki

1. はじめに

我が国では、1960年代に総理府統計局（当時）が空間統計を作成するための集計単位（グリッドまたはメッシュ）の定義について省庁横断的な研究を行い、地域メッシュコードと呼ばれる緯度と経度に基づく集計単位とコード付与体系を日本の標準として提唱した。さらに、このコード体系は日本工業規格として1974年に地域メッシュコード(JIS X0410)として標準化され現在に至っている[1]。これにより、政府統計を空間統計として公表する場合には、地域メッシュコードに従って公開することが現在に至るまで行われている。総務省統計局は e-Stat 内「地図で見る統計」を通じて国勢調査や経済センサスに基づく地域メッシュ統計を公開している[2]。また、国土交通省国土政策局国土政策課は国土交通省が有する国土数値情報や輸送、観光に関する位置情報を含む政府統計を地域メッシュ統計化して公開してきた[3]。

地域メッシュコードは、我が国で40年以上利用され続け、様々な種類の地域メッシュ統計が産業界および行政の分野で作成されている実績を有する。地域メッシュ統計は異なる統計を空間で連結し、統合分析を行うことができ、更に、任意の形状にメッシュ近似ではあるが再集計することが可能である。地域メッシュ統計は緯度経度を用いて統計を作成するため、行政区画の合併や変更など時間的な変化に対して影響を受けないという長所を有する。

測量法の「測量の基準」が日本測地系から世界標準である世界測地系(ISO 6709)に改正され、2002年4月1日から施行されたことにより、基準測量や公共測量は世界測地系で表示されなければならなくなった。これを受け、2002年に JIS X 0410 についても、10年間の移行期間を設けて1974年の地域メッシュコードの日本工業規格での定義を世界測地系に改め(JIS X 0410:2002)、2012年から世界測地系による地域メッシュコードの定義のみが標準とされるようになった。そのため、現在は世界測地系に従った地域メッシュ統計のみが公表されるようになっている。

このような背景から、現在我が国で研究され利用されてきた地域メッシュコード(JIS X0410)は日本以外の他国においても利用可能な状況となっている。

本稿では、この地域メッシュ統計を世界メッシュ統計として拡張し、これを全世界レベルで可視化・定量化することができる多言語の分析基盤について報告するとともに、開発方法と開発中の世界メッシュ統計分析基盤で利用可能な機能について一部紹介する。

2. 世界メッシュコードとは

発表者（佐藤・椿）は、世界測地系に従う地域メッシュコードの上位に2桁のコードを付加することによって、世界メッシュコード体系を構成できることを示した。この世界メッシュコードは地域メッシュコード(JIS X0410)の上位

互換性を有する全世界を覆うことができるメッシュコード体系[4]であり、図1に示すように、JIS X0410と同様に1次メッシュコードから6次メッシュコードまでの6階層の階層性を有し、緯度と経度から一意にメッシュコードが決まる。

特に、3次メッシュコード(10桁)は緯度 latitude と経度 longitude を与えた場合(1)式で定義され、1つの3次メッシュは互いにオーバーラップせず、緯度経度からのみから決定される緯度方向30秒、経度方向45秒の矩形集計単位を与える。

$$\text{3rd grid square code} = \begin{cases} o00p0uqvrw & (p < 10, u < 10) \\ o0p0uqvrw & (10 \leq p < 100, u < 10) \\ op0uqvrw & (p \geq 100, u < 10) \\ o00puqvrw & (p < 10, u \geq 10) \\ o0puqvrw & (10 \leq p < 100, u \geq 10) \\ opuqvrw & (p \geq 100, u \geq 10) \end{cases}$$

$$\begin{aligned} o &= 2^2x + 2y + 1 \\ (1 \leq o \leq 8, x = 0 \text{ or } 1, y = 0 \text{ or } 1, z = 0 \text{ or } 1) \\ p &= \lfloor (1 - 2x) \text{latitude} \times 60 \div 40 \rfloor \\ a &= \lfloor ((1 - 2x) \text{latitude} \times 60 \div 40 - p) \times 40 \rfloor \\ q &= \lfloor a \div 5 \rfloor \\ b &= (a + 5 - q) \times 5 \\ r &= \lfloor b \times 60 \div 30 \rfloor \\ c &= (b \times 60 \div 30 - r) \times 30 \\ u &= \lfloor (1 - 2y) \text{longitude} - 100z \rfloor \\ f &= (1 - 2y) \text{longitude} - 100z - u \\ v &= \lfloor f \times 60 \div 7.5 - v \rfloor \\ g &= (f \times 60 \div 7.5 - v) \times 7.5 \\ w &= \lfloor g \times 60 \div 45 \rfloor \\ h &= (g \times 60 \div 45 - w) \times 45 \end{aligned}$$

逆に3次世界メッシュコード *opuqvrw* が与えられている場合、そのメッシュの北西の緯度(latitude)と経度(longitude)は以下の式で与えられる。

$$\begin{cases} \text{latitude} = (1 - 2x) \times (p \times 40 \div 60 + q \times 5 \div 60 + (r - x + 1) \times 30 \div 3600) \\ \text{longitude} = (1 - 2y) \times (100z + u + v \times 7.5 \div 60 + (w + y) \times 45 \div 3600) \end{cases}$$

世界メッシュコードは地域メッシュコード(JIS X0410)に対して上位互換性を有するため、地域メッシュコード(JIS X0410)の上位に20を追加することにより、現状で国内において整備されている地域メッシュ統計をそのまま世界メッシュ統計として取り込むことが可能である。

総務省統計局が公開する我が国の市区町村別 3 次メッシュコード一覧[5]を参考として、GADM の Shape ファイルから世界 252 の国と地域に対する行政区世界メッシュデータの作成を行いこれの公開を行っている[6]。更に、JAXA だいち(ALOS) 30m 標高データに基づく、世界メッシュ標高統計、NASA 夜間光データに基づく世界メッシュ夜間光データの計算とその利用を行っている。これらのデータは世界メッシュ研究所[6]の Web ページから公開している。

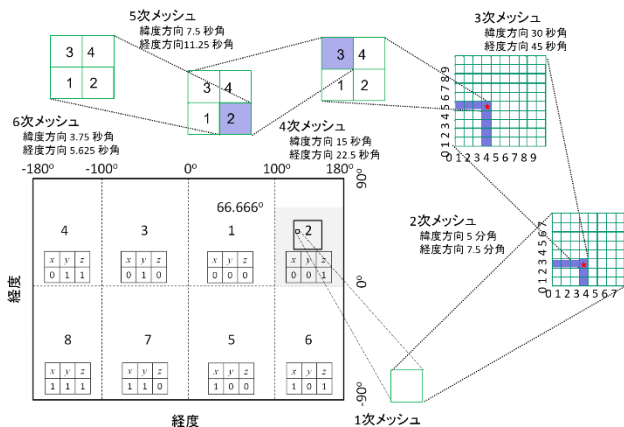


図 1 世界メッシュコード体系の概念図。1 次メッシュ：6 桁(緯度方向 40 分、経度方向 1 度)、2 次メッシュ：8 桁(緯度方向 5 分、経度方向 7.5 分)、3 次メッシュ：10 桁(緯度方向 30 秒、経度方向 45 秒)、4 次メッシュ：11 桁(緯度方向 15 秒、経度方向 22.5 秒)、5 次メッシュ：12 桁(緯度方向 7.5 秒、経度方向 11.25 秒)、6 次メッシュ：13 桁(緯度方向 3.75 秒、経度方向 5.625 秒)

3. 世界メッシュ統計データ可視化・分析基盤

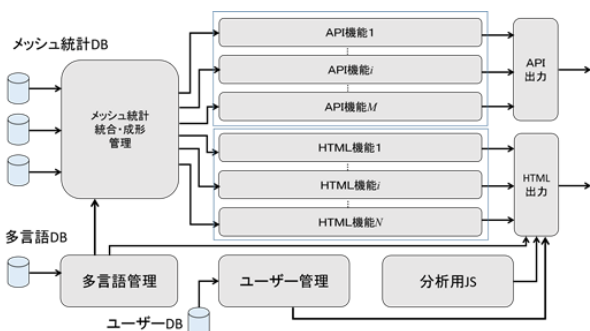


図 2 世界メッシュ統計データ可視化・分析基盤の概念図

本節では現在開発中の世界メッシュデータ可視化・分析基盤[4]について紹介する。この可視化・分析基盤はクラウドコンピュータ上に構築しており、クラウドコンピュータ上に構築した仮想サーバー(データベースサーバー、Web サーバー、データ収集サーバー)を用いて分析基盤の構築を行っている。HTML 形式による Web ブラウザでの閲覧を前提とした N 種類の機能と機械判読可能な形で出力を得る

M 種類の API 出力機能を有している。この世界メッシュ統計データ分析基盤の概念図を図 2 に示す。

世界メッシュ統計データは個別にデータベースとして格納されており、更に、国ごとに異なるテーブルとして保管されている。世界メッシュ統計統合・成形管理部はこれらの世界メッシュ統計データに対応する機能と選択された場所と領域(緯度経度および区画の大きさ)に応じて選別し統合してデータ配列として取り出す。HTML 出力の場合、これらを統合、成形したメッシュ統計データを機能ごとに峻別して出力する。API 出力の場合は指定された API 形式出力形式に変換後、機能ごとに峻別して機械判読可能な形で世界メッシュ統計データを出力する。

ユーザー管理部はユーザーデータベースを有しており、また、言語管理部は多言語テーブルを有している。ユーザー管理部に保管されるユーザー情報を参照して多言語テーブルで選択言語に世界メッシュ統計データのフィールド名を変換して世界メッシュ統計は読み出しされる。更に、HTML 形式出力に付加される文章についても、ユーザーの言語設定に応じて言語変換が行われて出力が行われる。多言語変換テーブルの作成は、Microsoft 社 Azure が提供する機械翻訳機能 API を使用し、英語と日本語を手作業により作成した後、これら 2 つの言語を基に多言語への機械翻訳を行い言語変換テーブルの作成を行った。その後、翻訳された各用語に対して、多言語テーブル内の機械翻訳結果を母語とし、英語または日本語を理解することができるユーザーに翻訳結果の確認修正作業を依頼することで、多言語テーブルの整備をすすめている。現在 9 か国語(英語、日本語、ドイツ語、イタリア語、スペイン語、韓国語、ベトナム語、中国語、ポーランド語)をサポートしている。

グローバル・システムの持続可能性評価基盤

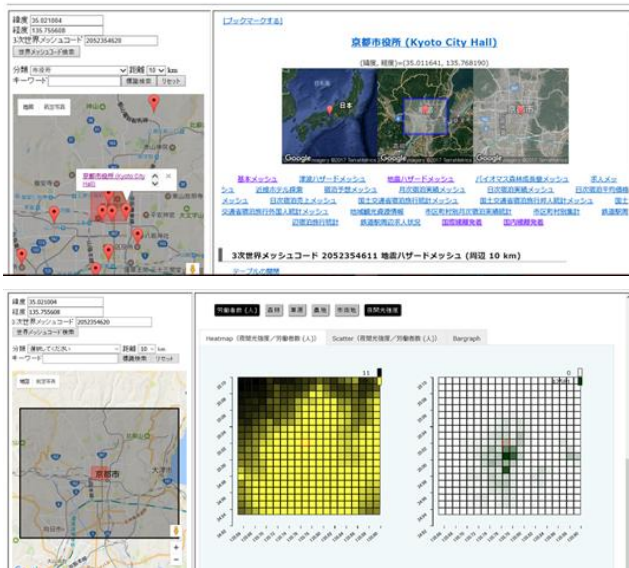


図 3 世界メッシュ統計可視化・分析基盤のスナップショット。(上) 位置選択画面の例。(下) 世界メッシュ統計の可視化の例。(左: NASA2012 年や夜間光強度 3 次メッシュ、右: 2012 年総務省統計局経済センサスに基づく労働者数 3 次メッシュ)

図 3 は図 2 で示した世界メッシュ統計データ可視化分析基盤のスナップショットである。左側に存在する、地図を用いて世界メッシュコードを選択（または、対象とする位置座標をカテゴリごとに選択）し、その周辺距離を指定することにより、矩形状に複数の世界メッシュデータをデータベースから抜き出す。抜き出された世界メッシュデータを空間上にヒートマップとして表示することにより、可視化し、かつ、2 種類の異なる世界メッシュデータに対する散布図の作成、線形回帰分析の機能を有する。機能ごとに選り出される世界メッシュデータの組み合わせを変え、または、指定された領域で再集計することにより希望する統計量の算出を行える機能を有する。

図 4 に、世界メッシュ統計データ可視化・分析基盤の状態推移図を示す。本システムはユーザーアカウントを作成する方式であり、ログイン画面からパスワード認証後トップ画面に推移する。トップ画面ではユーザー情報の更新、パスワード変更、通貨交換レートテーブルの表示、ログアウトが選択可能である。更に、トップページから場所選択画面に移動後、地図上から世界メッシュ統計をデータベースから抜き出すための位置とその周辺区画の選択を行いデータの抜き出しを行う。最初に表示されるのは基本世界メッシュ統計であり、その後、各機能に特化した世界メッシュ統計やそこから計算される時系列データの表示分析機能ごとに分岐できる。ユーザーごとに利用できる機能の制限を行うことができ、ユーザーは許可された機能のみを選び出し利用することができる。

現在の実装では、政府統計に基づくメッシュ統計、衛星データに基づくメッシュ統計及びインターネット上のホテル予約サイト・求人紹介サイトから収集したポイントデータに基づくメッシュ統計を世界メッシュ統計化してデータベース上に保管し、機能ごとにそれらの組み合わせを設計している。これら世界メッシュ統計は分析基盤を通じて Web ブラウザを通じて出力され、分析用 Javascript と対応させることで、Web ブラウザが動作する計算機上で可視化・解析を行い複数の統計表示機能を実装している。

世界メッシュ統計を作成するためのデータ源として、3 種類のデータ(政府統計に基づくメッシュ統計、衛星データに基づくメッシュ統計及びインターネット上のサイトから収集した位置情報付きデータ)に基づくメッシュ統計を利用しており、源データから、世界メッシュ統計を作成する方法として以下 4 種類の方法が存在する。

1. 政府統計として公開される地域メッシュ統計を世界メッシュ統計として変換する。国内で流通する地域メッシュコードの上位に 20 を付加する。または他国の異なる集計単位で作成されたグリッド統計を面積案分計算により世界メッシュ統計により近似する。

2. 世界メッシュコードで定義される区画より小さな区画で集計された統計または観測データ（4 次地域メッシュデータより小さな地域メッシュデータおよび衛星データ）からより大きなメッシュ統計を作成する。

3. 緯度経度を含むポイントデータから世界メッシュコードで定義される区画を利用して集計を行う。

4. 多角形（ポリゴン）として表現されるラベル付き閉領域を世界メッシュコードで表現される区画が含まれるかを検査することによりラベルに関する世界メッシュデータを作成する。

更に、現在開発中の定量化・可視化機能は大別すると 6 種類存在している。

- (1) 世界メッシュ統計を視覚的に表示する機能
- (2) 2 種類の世界メッシュ統計を散布図として表示して回帰分析を行う機能
- (3) ある世界メッシュに対して存在する値を一覧として表示する機能
- (4) 日次ごとの世界メッシュ統計を選択範囲で集計して、時系列データとして表示する機能
- (5) 世界メッシュ統計を行政界メッシュデータに従い集計することにより、行政区画の統計に再集計する機能
- (6) 世界メッシュ統計を選択された区画で再集計することにより、選択された位置の周辺の統計を作成する機能である。

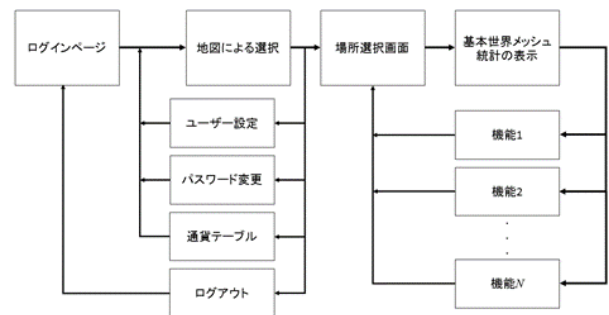


図 4 世界メッシュ統計可視化・分析基盤の状態推移図

4. 世界メッシュ統計可視化・分析基盤の開発手法

ビッグデータ標準化の文脈では、ビッグデータ参照モデルと呼ばれる役割分担とレイアーを規定したモデルが NIST により提案された。この、ビッグデータ参照モデルでは、ビッグデータ利活用で現れる役割とその名称およびそれらの関係を定義している[8]。このビッグデータ参照モデルの概念図を図 5 に示す。

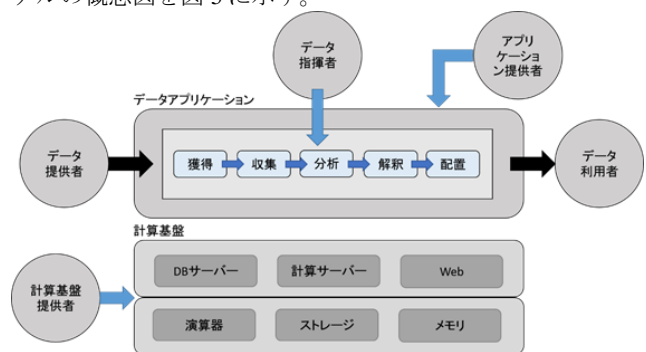


図 5 ビッグデータ参照モデル

このビッグデータ参照モデルには 5 種類の役割（データ利用者、データ提供者、データ指揮者、データアプリケーション開発者、データ基盤運営者）が定義されており、データ利用者とデータ提供者を結びつける形としてデータアプリケーションが存在しており、データアプリケーションはデータ指揮者の管理のもと開発・運用がなされ、これら

はデータ基盤運営者が運営する計算資源の上で稼働するというモデルである。

本稿で紹介した世界メッシュ統計可視化・分析基盤はデータアプリケーションとして位置付けられるため、データ提供者とデータ利用者を結合するアプリケーションである。データアプリケーションを開発する場合、データ利用者とデータ提供者とを相互に結びつけるためにデータ利用者にとって有益となるアプリケーションを開発する必要があるが、一般にこの問題はチキン・エッグ問題に直面する。

データ提供者はニーズに基づきデータを作成しデータを提供したいと考える。この活動はデータ利活用の事例が存在している必要があるが、アプリケーションとデータ利用者が存在していない状況ではデータ提供者が提供するデータを決められないという問題がある。他方、データ利用者はアプリケーション内に希望のデータが存在していれば、利用したいと考えているが、具体的なデータの利用事例についてはアプリケーションが見えるまで理解できない。そのため、どのようなデータアプリケーションを開発すればよいかデータアプリケーション開発者はデータ提供者とデータ利用者双方に対して事例を提示する必要がある。

このチキン・エッグ問題を解決するために、本稿ではデータ提供者として既にオープンデータとしてデータを提供している行政組織（総務省統計局、統計センター、国土交通省）が公開する地域メッシュ統計とフリーのインターネットデータ API、および、JAXA と NASA が公開する衛星画像、衛星データをデータ源として世界メッシュ統計を構築した。他方、これらのオープンデータから作成される複数の世界メッシュ統計を基に、データ利活用の事例研究をデータ利用者となる可能性のある組織の実務家または研究者と共同研究を実施することにより集め、それら事例研究に基づきアプリケーション開発を行った。

この時、他の組織の実務内容についてエスノグラフィーの手法（共同研究に基づく聞き取りと参与調査）を繰り返し利用することでデータ利活用事例の掘り起こしを行い、アジャイル開発の手法を用いてアプリケーションの開発を進めた。このエスノグラフィーの手法を用いた分析は、ビッグデータ参照モデル上では、データ指揮者が行う役割に相当する。

アジャイル開発とはソフトウェア開発方法論の一種であり、ソフトウェアの仕様を厳密に決定してから、コーディング作業に入るウォーターフォール型開発とは異なり、小規模なソフトウェア開発を行った後、ソフトウェアの検証作業を実フィールドで行い、その時のソフトウェア利用に関するユーザーの意見を取り入れることにより、次のソフトウェア開発の仕様とする操作を繰り返す PDCA サイクルを取り入れたソフトウェア開発の手法である。本稿で紹介した世界メッシュ統計分析基盤のアジャイル開発のために、執筆現在(2017年6月)約100名のデータ提供者とデータ利用者が本プロジェクトに協力している。

5. 世界メッシュ統計利活用の事例

地域メッシュ統計の利活用ユースケースとしては以下の6種類が想定されている[9]。

(1) データ・リンケージとデータの加工: 地域メッシュ単位のデータを相互に結合（リンケージ）して利用すること

が可能である。データの演算や加工により新しいデータの合成が可能である。

(2) 地域メッシュ地図: 地域メッシュデータを地図上に表示することにより、ある地域の構造を視覚的にとらえることが可能である。更に、多数のデータについて地域メッシュ統計を用いることで分析が可能となる。

(3) 任意の地域についてのデータ作成: ある与えられた区画を表現する地域メッシュ統計を合算または平均することにより、必要なデータを再集計により得ることが可能となる。

(4) 圏域の決定: 地域メッシュ間の距離が簡単に求められることを利用して、距離により与えられる圏域内の活動や、需要を評価し、最適配置の問題に適用することが可能である。

(5) 観察単位としての地域メッシュ: 観察単位としてみることにより、地域メッシュ内のデータを相互に蓄積し、分析することが可能である。

(6) シミュレーションの単位としての地域メッシュ: 地域メッシュをシミュレーションの単位とすることで、パーコレーションモデルによる拡散シミュレーションや、移動シミュレーションに地域メッシュを利用することが可能である。

地域メッシュ統計を用いた経済社会活動の定量的評価に関する研究[10,11,12]や、災害リスクの見積もり[13,14]について様々な先行研究がある。例えば、地震ハザードの地域メッシュデータについては、防災科学技術研究所よりデータが公開されている[14]。

ここでは、水害被害想定に関する地域メッシュデータをどのように作成したかの経緯と、この水害被害想定メッシュデータを他のデータとリンケージすることにより得られた水害被害想定に関する見積もり結果について報告する。

水害の被害想定地域メッシュデータは現時点では公開されているものは存在していないため、国土交通省国土政策局国土情報課が提供する平成24年度国土数値情報浸水想定区画データのポリゴンデータ[15]をもとに、3次地域メッシュに変換を行い浸水深クラスの3次地域メッシュデータを作成した。

浸水想定区画データは河川管理者（国土交通大臣、都道府県知事）から提供された浸水想定区域図をもとに都道府県ごとにポリゴンデータとして整備したものである。浸水想定区域データのもととなる浸水想定区域図は水防法(昭和24年法律第193号)に基づいて整備されている。水防法第十条第二項及び第十一条第一項に基づき指定される洪水予報下線並びに水防法第十三条に基づき指定される水位周知河川の内、各河川管理者より資料提供を受けられたものを基に作成されている。各河川管理者が作成した浸水想定区域図のGISデータや数値地図データ、浸水想定区域図の画像データ、紙の浸水想定区域図をスキャンにより電子化した画像データから、浸水深ごとにポリゴン（面）形式で表現された、シェープファイルが都道府県別にそれぞれ公開されている。このポリゴンデータを基に、地域メッシュ統計化の作業を行うことにより浸水想定区域3次地域メッシュデータの作成を行った。

ポリゴンデータからメッシュデータへの変換方法は以下のとおりである。浸水想定区画は浸水深ごとに異なるポリゴンとして保管されている。ポリゴンとして表現される閉

曲線領域を包含する矩形領域を特定した後、その矩形領域内に存在する全ての 3 次メッシュに対して閉曲線領域を含むか、含まないかを評価する。検査した 3 次メッシュが閉曲線内の領域を含む場合には、地域メッシュコードと閉曲線が対応する浸水深クラスを出力する。この検査を、該当する 3 次メッシュ全てに対して行う。図 5 は想定浸水深レベルの 3 次メッシュデータである。色の濃淡が浸水深レベルに対応している。



図 5 平成 24 年度国土交通省水深想定区域浸水深 3 次メッシュデータの可視化図。色の濃淡が浸水深レベルに対応する。

ここで抽出された浸水想定区画浸水深 3 次メッシュデータと 2010 年総務省統計局国勢調査 3 次メッシュ統計[16]、および、2012 年総務省統計局経済センサス労働者、事業者数 3 次メッシュ統計[17]とを用いることにより、浸水深想定区画内に暮らす各浸水深レベルにおける人口、労働者数、事業所数の概算を行った。計算結果を表 1 に示す。この表から 5.0m 以上の浸水深クラスを有する 3 次メッシュに 88.5 万人(33.8 万世帯)が暮らしており、39,142 事業所で働く 38.0 万人の労働者が存在していることが分かる。また、2.0-5.0m の想定浸水深にクラス人口は 1092.5 万人(453.4 万世帯)が暮らしており、54.3 万事業所で 547.6 万人の労働者が働いている。更に、1.0m 以上の浸水深を有するメッシュに暮らす人口は、2322.9 万人であり、これは 2010 年国勢調査で確認される日本の総人口の 18.1%に相当する。

表 1 浸水深クラスごとの人口、事業所数、労働者数の集計結果

浸水深レベル	3 次メッシュ数	人口(人)	事業所数	労働者(人)
0-0.5m 未満	5,617	11,261,338	629,893	6,976,519
0.5-1.0m 未満	3,565	6,166,131	328,678	3,466,063
1.0-2.0m 未満	6,155	11,418,739	603,637	5,979,277
2.0-5.0m 未満	6,889	10,925,326	543,858	5,476,574
5.0m 以上	1,116	885,905	39,142	380,971

本稿で提案する世界メッシュ統計データ分析基盤を用いることにより場所ごとの想定浸水深と政府統計による国

勢調査を連動させてインタラクティブに空間リスクの計量を行うことが可能である。

地域メッシュ統計から再集計を行う場合、再集計区画面積に依存して、インタラクティブな計算を行うことが困難となる場面がある。特に、国レベルでの再集計作業がそれに該当する。このような場合、事前に再集計計算をバッチ処理的に行い計算結果に関するデータベースをシステム内に保有しておく必要がある。

6. LOD による世界メッシュ統計データの作成

統計センターが運用している e-Stat では、オープンデータの最上位レベルと言われている Linked Open Data(LOD)で統計データの提供を 2016 年から開始をしている[18]。

LOD は World Wide Web Consortium(W3C)から勧告されている標準的な技術で構成されており、各種データは Resource Description Framework(RDF)を用いて作成し、データの検索には SPARQL Query Language(SPARQL)を利用する。また、統計データなどの多次元データを RDF 化する方法としては、RDF Data Cube Vocabulary が W3C から勧告されており、e-Stat ではこれを用いて作成している。

RDF Data Cube Vocabulary では、統計データを次元(Dimension)、測度(Measure)、観測値(Observation)、属性(Attribute)の 4 つの要素で構成する。次元は、統計データを何で分類したかを表すもので、地域、時間、性別、年齢などが該当する。測度は、統計データを何で集計等したものか表したもので、人、価格、指数などが該当する。観測値は、統計データの数値そのものを表すものである。属性は、統計データに付随する属性を表すもので、単位、状態などを表すものである。実際の統計表で説明すると図 6 のようになる。

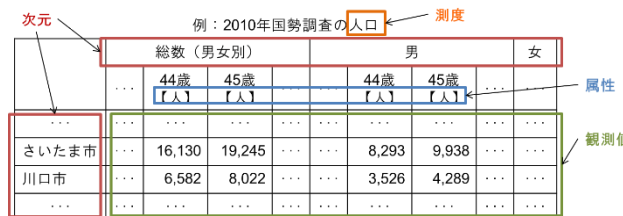


図 6 統計 LOD における定義

上記の概念を基に、1 つの統計データ(セル)ごとに URI を付与し、次元などの定義とリンクする形で統計データを RDF に変換し、国勢調査などのデータの一部を公開している。

現時点では、メッシュ統計データは LOD として公開していないが、上記の方法を用いて国勢調査などのメッシュ統計データを提供することを予定している。国勢調査のメッシュ統計は、性別人口などを 3 次、4 次、5 次メッシュで公開しており、メッシュコードは地域を表す属性であるため、次元の一つとして各データとリンクすることとなる。また、それぞれのメッシュコードに世界メッシュコードを用いて URI を定義し、基本情報(緯度経度情報など)をリンクすることを考えており、イメージは図 7 の通りとなる。

メッシュ統計データを世界メッシュコードとともに RDF 化し LOD として公開することにより、統計データの所在が明確化され、データ定義も標準化される。さらに、世界メ

ッシュコードが URI として定義されるため、統計データ以外の他のデータとのリンクが容易となり、世界メッシュコードをキーとしたデータの取得が可能となることが期待される。

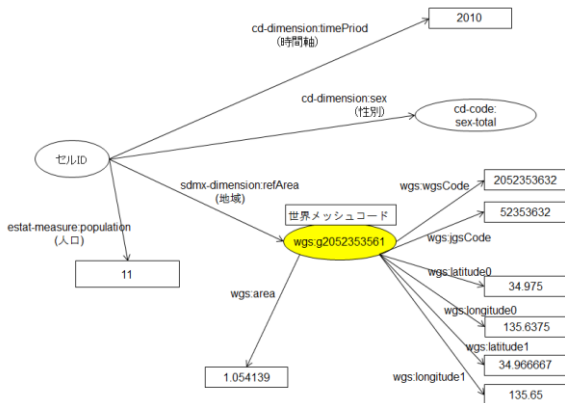


図 7 メッシュ統計データの RDF 化

7. まとめと今後の課題

本稿では、空間と紐づけることによりデータの連結を可能とする可視化定量システムの提案を、世界メッシュコードと世界メッシュ統計により実現する方法および、そのような可視化定量システムが有する利活用の可能性について言及した。世界メッシュコード体系について概要を述べるとともに、政府統計から公開されているデータの例、分析基盤の構造について示した。本稿では、水害被害想定メッシュデータをポリゴンデータから作成することにより、総務省統計局国勢調査、経済センサス地域メッシュ統計と結合分析を行い、浸水深クラスごとに存在する経済社会的価値について見積もりを行える事例を示した。また、オープンデータでの活用が期待されている LOD を用いたメッシュデータの作成についても触れ今後の発展性について紹介した。

一般に、データアプリケーションの開発においては、データ提供者とデータ利用者をどのように増やし、データ利活用の事例およびユースケースをどのように開発していけばよいかという問題がある。

本稿で述べた世界メッシュ統計データ可視化・分析基盤の開発では、データ提供者としてオープンデータを提供する政府組織および企業に協力を求めつつ、データ利用者となり得る潜在的な組織内の実務家・研究者との共同研究を通じてデータ利活用の事例とユースケースを獲得しながらすすめている。特に、データ分析基盤の開発をアジャイル開発により PDCA サイクル的に発展させることによりアプリケーションの機能を充実させながら、データ提供者とデータ利用者を増やすことでコミュニティーの形成をはかる手法を採用した。

今後の課題として、世界メッシュ統計データの拡充のためのデータ提供者の組織化による世界メッシュ統計の拡充およびそのための世界メッシュコード計算ライブラリの普及が挙げられる。また、世界メッシュ統計データを利用する利用者の組織化とユースケース開発、および世界メッシュ統計データ分析基盤へのアジャイル開発（実装と実環境中での評価）の継続が挙げられる。

謝辞

本研究は科学技術振興機構(JST)戦略的創造研究推進事業(さきがけ)「ビッグデータ統合利活用のための次世代基盤技術の創出・体系化」(研究統括:喜連川優、副統括:柴山悦哉)のもと「グローバル・システムの持続可能性評価基盤に関する研究」(研究代表者:佐藤彰洋)の資金に基づき実施されています(Grant Number: JPMJPR1504; 研究期間:2015年10月-2019年3月)。本研究で示した世界メッシュ統計および世界メッシュデータは2016年度京を中核とする HPCI システム課題研究 (hp160060) 「経済社会環境データを用いた経済社会システムのストックとフローに対する大規模計算」(計算資源提供:統計数理研究所)の計算資源を用いて計算されました。

参考文献

- [1] 国土実態総合統計の開発・整備に関する研究報告(昭和46年)発行:総理府統計局
- [2] 総務省統計局, e-Stat 地図で見る統計 GIS, <http://www.e-stat.go.jp/SG1/estat/toukeiChiri.do?method=init>
- [3] 国土交通省国土政策局国土情報課, 国土数値情報ダウンロードサービス, <http://nlftp.mlit.go.jp/ksj/index.html>
- [4] 佐藤彰洋, 椿広計, "ビッグデータ時代に必要な標準化", 統計, 2015年9月号(2015) pp. 32-38.
- [5] 総務省統計局市区町村別メッシュコード一覧, http://www.stat.go.jp/data/mesh/m_itiran.htm
- [6] 世界メッシュ研究所, <http://www.fitsus.jp/worldgrids/>
- [7] MESHSTATS, <https://www.meshstats.xyz/>
- [8] 佐藤彰洋, "ビッグデータと標準化", 標準化と品質管理, Vol. 70, No. 3 (2017) pp.20-23.
- [9] 伊藤彰彦, 地域メッシュ統計の紹介, 行動計量学 Vol. 4, No. 1 (1976) pp. 59-63.
- [10] Aki-Hiro Sato, Tsutomu Watanabe, "Measuring Activities and Values of Industrial Clusters based on Job Opportunity Data Collected from an Internet Japanese Job Matching Site", 2016 IEEE International Conference on Big Data (Big Data), 5-8 Dec. 2016, 2016, pp. 2199-2208
- [11] Aki-Hiro Sato, Chihiro Shimizu, Takayuki Mizuno, Takaaki Ohnishi, Tsutomu Watanabe, "Relationship between job opportunities and economic environments measured from data in internet job searching sites", Procedia Computer Science, Volume 60 (2015) pp. 1255-1262
- [12] Aki-Hiro Sato, "Microdata analysis of the accommodation survey in Japanese tourism statistics", 2015 IEEE International Conference on Big Data (Big Data), Oct. 29 2015-Nov. 1, 2015, pp. 2700-2708.
- [13] Aki-Hiro Sato, Hidefumi Sawai, "Geographical risk assessment from tsunami run-up events based on socioeconomic-environmental data and its application to Japanese air transportation", Procedia CIRP, Volume 19 (2014) pp. 27-32.
- [14] 防災科学技術研究所, 地震ハザードステーション, <http://www.j-shis.bosai.go.jp/>
- [15] 国土交通省国土政策局国土情報課 国土数値情報 浸水想定区域データ(平成24年度), <http://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-A31.html>
- [16] 総務省統計局 2010年度国勢調査3次メッシュデータ, 地図で見る統計(統計GIS), <http://e-stat.go.jp/SG2/eStatGIS/page/download.html>
- [17] 総務省統計局 2012年経済センサス3次メッシュデータ, 地図で見る統計(統計GIS), <http://e-stat.go.jp/SG2/eStatGIS/page/download.html>
- [18] 政府統計の総合窓口(e-Stat)-統計 LOD, <http://data.e-stat.go.jp/lodw/>

† 京都大学 Kyoto University

‡ 独立行政法人統計センター National Statistics Center